

HABILITATION   DIRIGER DES RECHERCHES (HDR)

DE
L'UNIVERSIT  D'ANGERS

Sp cialit  : *Informatique*

Par

Nicolas Gutowski

**Apprentissage automatique pour la g n ration et l'exploration
Adaptation et contextualisation pour donn es complexes**

Travaux pr sent s et soutenus   Angers, le 31/03/2025

Unit  de recherche : LERIA (EA 2645)

Rapporteurs avant soutenance :

Thi-Bich-Hanh Dao, Ma trese de conf rences (HDR), Universit  d'Orl ans (LIFO), Orl ans

Allel Hadjali, Professeur des Universit s, ENSMA (LIAS), Poitiers

Arnaud Liefoghe, Professeur des Universit s, Universit  du Littoral C te d'Opale (LISIC), Calais

Composition du jury :

Pr sident : Fr d ric Koriche, Professeur des Universit s, Universit  d'Artois (CRIL), Lens

Examineurs : Thi-Bich-Hanh Dao, Ma trese de conf rences (HDR), Universit  d'Orl ans (LIFO), Orl ans

Allel Hadjali, Professeur des Universit s, ENSMA (LIAS), Poitiers

Arnaud Liefoghe, Professeur des Universit s, Universit  du Littoral C te d'Opale (LISIC), Calais

Directeur de recherche : Fr d ric Saubion, Professeur des Universit s, Universit  d'Angers (LERIA), Angers

REMERCIEMENTS

« *We can only see a short distance ahead, but we can see plenty there that needs to be done.* » [Tur09] **Alan Turing - MIND, 1950**

Je tiens à exprimer ma sincère gratitude à toutes les personnes qui, de près ou de loin, ont contribué à mon parcours de recherche. Leur soutien précieux m'a permis de réaliser ce travail d'Habilitation à Diriger des Recherches dans des conditions favorables. Comme l'a si bien exprimé Alan Turing, même si notre regard ne porte que sur une courte distance, il reste encore tant à accomplir. C'est dans cet esprit d'exploration continue que je souhaite poursuivre mes recherches, conscient du chemin déjà parcouru et des défis qui restent à relever.

Merci à mes collègues chercheurs, Tassadit Amghar, Sébastien Aubin, Olivier Camp, Fabien Chhel, Benoit Da Mota et Daniel Schang, pour leur précieuse collaboration dans divers projets de recherche. Merci à Jean-Pierre Briot, Amal El Fallah Seghrouchni, Sylvain Lamprier et Philippe Martinet pour leurs orientations scientifiques avisées dans le cadre des différentes directions de thèse qu'ils ont portés. Merci à Thomas Cauchy et Pierre Abraham experts des domaines de la chimie et du médical, pour avoir pris de leur temps pour m'expliquer certains détails de leur discipline respective.

Merci aux doctorants et post-doctorants que j'encadre ou que j'ai encadrés : Alexandre Letard (2020), Nathan Fradet (2021), Fabien Lioni (2022), Derin Özer (2023), Gaëlle Milon-Harnois (2024, post-doctorante) pour leur énergie et pour porter ou avoir su porter nos travaux scientifiques avec brio.

Merci aux rapporteurs de mon HDR : Allel Hadjali (LIAS), Arnaud Liefoghe (LISIC) et Thi-Bich-Hanh Dao (LIFO) pour avoir accepté d'évaluer mon travail ainsi qu'à Frédéric Koriche (CRIL) pour être membre et président du jury.

Merci à Alain Clément, Chef du Département Génie Électrique et Informatique & Industrielle, à Patrice Mangeard Directeur de l'IUT d'Angers et à Éric Monfroy Directeur du LERIA qui ont su m'aménager les conditions de travail adéquates pour la réalisation de mon HDR. Un grand merci à Frédéric Saubion, Directeur de recherche de mon HDR, qui tout en se portant garant m'a accompagné et m'a conseillé.

Merci également aux diverses entreprises et institutions pour leur financement : Kara Technology pour la thèse CIFRE d'Alexandre Letard, Aubay pour la thèse CIFRE de Nathan Fradet, la Direction Générale de l'Armement et l'INRIA pour le financement de la thèse AID de Fabien Lioni, l'Université d'Angers pour la thèse de Derin Özer et enfin la région Pays de la Loire pour le financement du projet ^GGENOCOD dans le cadre du dispositif « Étoiles Montantes ».

Enfin, je tiens tout particulièrement à remercier ma femme et mes fils pour leur présence, leur soutien et leur amour.

SOMMAIRE

Introduction	11
Adaptation et contextualisation dans l'apprentissage automatique pour données complexes	11
Organisation du mémoire	13
Présentation synthétique	17
Résumé du profil	17
Curriculum Vitæ	18
Activités d'enseignement et charges administratives associées	20
Activités de recherche	22
Rayonnement, animation et responsabilités scientifiques	33
Encadrements	36
Production scientifique	37
Réflexion prospective	40
I Optimisation multicritère appliquée à l'apprentissage	43
1 Apprentissage par renforcement et systèmes de recommandation multicritères	46
1.1 Concepts généraux sur les systèmes de recommandation et les algorithmes de bandits-manchots	48
1.1.1 Rappel sur les systèmes de recommandation	48
1.1.2 Rappel sur les bandits-manchots pour la recommandation	49
1.1.2.1 Définitions	49
1.1.2.2 Bandits-manchots pour la recommandation : Définitions	50
1.1.2.3 Bandits-manchots contextuels pour la recommandation : Définitions	51
1.1.2.4 Limites des <i>MAB/CMAB</i>	52
1.1.3 Rappel sur bandits-manchots pour la recommandation de plusieurs éléments	53
1.1.3.1 Bandits-manchots pour la recommandation de plusieurs éléments : Définitions	53
1.1.3.2 Fonctions de récompenses pour bandits-manchots à tirage multiple	54
1.2 Positionnement des contributions	55
1.3 Méta-sélection d'algorithmes d'apprentissage	57

1.3.1	Motivation	57
1.3.2	Concept de la sélection automatique d'algorithmes	57
1.3.3	Cas de la sélection automatique d'algorithmes pour la recommandation	58
1.4	Méthodes <i>Gorthaur</i> : Méta-apprentissage pour la recommandation	59
1.4.1	Cœur des méthodes <i>Gorthaur</i> : le Compass	60
1.4.2	Méthodes <i>Gorthaur</i> à tirage simple	61
1.4.2.1	Fonctionnement des méthodes <i>Gorthaur</i> à tirage simple	61
1.4.3	Problème de bandit-manchot de sélection d'algorithmes multicritère : Définitions	64
1.4.3.1	Le portefeuille possible de <i>Gorthaur</i>	65
1.4.3.2	Les promesses des méthodes <i>Gorthaur</i>	65
1.4.3.3	Méthode <i>Gorthaur-EXP3</i>	66
1.4.4	Méthodes <i>Gorthaur</i> à tirage multiple	70
1.4.4.1	Sélection d'algorithmes pour la recommandation multicritère de type top-k	70
1.4.4.2	Nouvelles approches : <i>Multiple-Play Gorthaur</i> et <i>Budgeted Gorthaur</i>	71
1.4.4.3	Méthode <i>Multiple-play Gorthaur</i>	72
1.4.4.4	Méthode <i>Budgeted-Gorthaur</i>	76
1.4.5	Synthèse des résultats et discussion sur les méthodes <i>Gorthaur</i>	79
1.4.6	Réflexion critique	81
2	Optimisation multicritère pour la sélection d'attributs médicaux	82
2.1	Problème de sélection d'attributs multi-objectif	84
2.1.1	Rappel de l'état de l'art sur l'optimisation multi-objectif	85
2.1.2	Comparaison de méthodes de type métaheuristique populaires en sélection d'attributs multi-objectif.	88
2.1.2.1	Méthodes de sélection	88
2.1.2.2	Principales fonctions d'évaluation	88
2.1.2.3	Choix des méthodes comparées	89
2.1.2.4	Autres méthodes qui pourraient être considérées.	91
2.2	Positionnement de nos contributions	91
2.3	Genetic Algorithm with Multi-Objective Compass - <i>GAwC</i>	92
2.3.1	Formalisation du problème multi-objectifs	92
2.3.2	La méthode <i>GAwC</i>	93
2.3.2.1	Algorithme génétique original	93
2.3.2.2	Genèse et concept de <i>GAwC</i>	95
2.3.2.3	<i>GAwC</i> : Genetic Algorithm with Multi-Objective Compass	95
2.3.3	Synthèse des résultats et discussion sur la méthode <i>GAwC</i>	99
2.3.4	Réflexion critique	102
2.3.4.1	Rapprochement avec <i>MOEA/D</i>	102

2.3.4.2	Éthique des décisions algorithmiques.	103
---------	---	-----

II Apprentissage profond et données complexes pour l'estimation et la génération **105**

3	Évaluation de comportements dynamiques non linéaires : application aux véhicules terrestres	109
3.1	Apprentissage informé par la physique : modèles hybrides	110
3.1.1	Modèles d'apprentissage informés par la physique	111
3.1.2	Applications dans le domaine de la dynamique des véhicules	112
3.2	Positionnement de nos travaux de recherche	113
3.3	Estimation de paramètres des systèmes dynamiques non linéaires : application à l'évaluation de la dynamique latérale des véhicules	116
3.3.1	Modélisation de la dynamique latérale d'un véhicule	116
3.3.1.1	Modèle bicyclette à deux degrés de liberté	116
3.3.1.2	Modèle de Pacejka	117
3.3.2	Estimation bayésienne des paramètres de systèmes dynamiques non linéaires	118
3.3.2.1	Estimation des paramètres de systèmes dynamiques non linéaires : vers une approche bayésienne	119
3.3.2.2	Cadre bayésien	121
3.3.2.3	<i>Approximate Bayesian Computation (ABC)</i>	122
3.3.2.4	Notre méthode bayésienne pour l'estimation	123
3.3.2.5	Synthèse des résultats	125
3.3.2.6	Discussion	126
3.3.3	Estimation des paramètres de la dynamique des véhicules par apprentissage par transfert guidé par la physique	128
3.3.3.1	Estimation des paramètres de la dynamique des véhicules : vers une approche hybride	128
3.3.3.2	Notre méthode hybride en trois étapes	130
3.3.3.3	Synthèse des résultats	133
3.3.3.4	Discussion	135
3.4	Estimation d'état en temps réel par apprentissage profond et estimation à horizon glissant	136
3.4.1	Estimation d'état des systèmes dynamiques	136
3.4.2	Réseaux de neurones pour l'estimation d'état et la résolution de problèmes inverses	137
3.4.3	Méthode proposée	138
3.4.3.1	Les 3 points clés de notre contribution	138
3.4.3.2	Description générale de notre méthode	139
3.4.3.3	Description formelle et détaillée de notre méthode	140

3.4.4	Synthèse des résultats	142
3.4.5	Discussion	145
3.5	Réflexion critique	146
4	Optimisation des représentations symboliques musicales pour l'apprentissage profond	148
4.1	Concepts généraux sur la génération de musique symbolique par apprentissage profond	149
4.1.1	Rappel sur les représentations de la musique.	150
4.1.2	Formats de données pour la représentation et la génération de musique symbolique	150
4.1.2.1	Formats numériques	150
4.1.2.2	Cas spécifiques de la représentation sous forme de tokens	152
4.1.3	Tokenisation de la musique symbolique	153
4.1.3.1	Principe de la tokenisation	153
4.1.3.2	Principales méthodes de tokenisation	153
4.1.3.3	Enjeux de la tokenisation en musique symbolique	154
4.1.4	Modèles discrets et continus pour la génération de musique symbolique	155
4.1.5	Mécanisme d'attention et <i>Transformers</i>	156
4.1.5.1	Mécanisme d'attention	157
4.1.5.2	Architecture <i>Transformer</i>	157
4.2	Positionnement de nos travaux de recherche	159
4.3	Contribution logicielle : <i>Miditok</i>	160
4.3.1	Avant <i>Miditok</i>	160
4.3.2	<i>Miditok</i> : paquet Python pour la tokenisation des fichiers <i>MIDI</i>	161
4.3.2.1	Une contribution logicielle	161
4.3.2.2	Tokeniser avec <i>Miditok</i>	161
4.4	Impact de la tokenisation du temps et de la durée des notes sur la modélisation de la musique symbolique	162
4.4.1	Méthodologie	163
4.4.1.1	Modèles et entraînements	163
4.4.1.2	Caractéristiques techniques	163
4.4.1.3	Tokenisations	163
4.4.1.4	Tâches évaluées	164
4.4.2	Synthèse des résultats	166
4.4.3	Discussion	166
4.5	Byte Pair Encoding pour la musique symbolique	167
4.5.1	Contexte et motivation	167
4.5.2	Stratégies de réduction de la longueur des séquences	168
4.5.3	<i>Byte Pair Encoding (BPE)</i>	168
4.5.4	Synthèse des résultats	170

4.5.5	Synthèse et discussion	170
4.6	Réflexion critique générale sur les <i>LLMs</i>	171
III	Bilan et perspectives	173
5	Conclusion et perspectives de recherche	174
5.1	Bilan de mes recherches	174
5.2	Perspectives de recherche	176
5.2.1	Travaux en cours et perspectives liées au projet ^G <i>GENOCOD</i>	176
5.2.1.1	Contexte applicatif et genèse du projet	176
5.2.1.2	Objectifs du projet ^G <i>GENOCOD</i>	177
5.2.1.3	Correspondances avec les approches présentées dans les chapitres précédents	178
5.2.2	Genèse des <i>Work Packages</i> de ^G <i>GENOCOD</i>	179
5.2.2.1	Synthèse des <i>Work Packages</i> en cours du projet ^G <i>GENOCOD</i> et perspectives	181
5.2.2.2	Organisation du projet ^G <i>GENOCOD</i>	182
5.2.2.3	Impacts et retombées du projet	182
5.2.3	Perspectives générales et projets de recherche futurs	184
5.2.4	Conclusion	184
	Bibliographie	185
	Annexes	210
A	Contributions complémentaires et informations complémentaires sur les contributions.	210
A.1	Exploitation de retours utilisateur partiels	210
A.1.1	<i>BSB</i> et <i>BUSBC</i>	210
A.1.2	Synthèse des résultats obtenus	212
A.1.3	Synthèse et analyse réflexive sur les méthodes <i>P-BSB</i> et <i>BUSBC</i>	213
A.2	Exemple de <i>Budgeted-Gorthaur</i> avec <i>EXP3-754515</i>	213
A.3	Compléments d'information sur l'algorithme génétique utilisé dans la méthode <i>GAwC</i>	214
A.4	Quatre points de vue du <i>Compass</i> - Exemple concret	215
A.5	Tokenisation dans <i>MidiTok</i>	215
A.6	Tokens supplémentaires	216

B	Pseudo-codes d'algorithmes de l'État de l'art employés dans nos travaux	217
B.1	Algorithme ε -Greedy	217
B.2	Algorithme <i>UCB</i>	218
B.3	Algorithme <i>Thompson Sampling</i>	218
B.4	Algorithme <i>EXP3</i>	219
B.5	Algorithme <i>LinUCB</i>	219
B.6	Algorithme <i>Linear Thompson Sampling</i>	220
B.7	Algorithme <i>PV-BUSBC-OE</i>	220
C	Liens des pages web archivées	221

INTRODUCTION

Adaptation et contextualisation dans l'apprentissage automatique pour données complexes

L'Intelligence Artificielle (IA) et plus particulièrement l'apprentissage automatique, ont connu une accélération depuis les années 2010 portée entre autres par : l'explosion des performances par l'apprentissage profond (*Deep Learning*), l'accélération des calculs à l'aide de GPUs toujours plus puissant, la multiplication des dépôts publics (codes) sur des plates-formes type *GitHub* (<https://github.com/>). Cet essor a transformé radicalement notre capacité à traiter, analyser et générer des données plus ou moins « complexes » dans des domaines variés. Alors que le volume et la diversité des données ne cessent de croître selon le paradigme du « *Big Data* », les enjeux liés à l'adaptation, à la contextualisation, à la représentation, et à l'optimisation au regard de ces données, deviennent de plus en plus cruciaux. Comme l'a exprimé Judea Pearl, sans la capacité à extraire des relations causales des données, les systèmes d'IA seront condamnés à des approximations superficielles et à des décisions peu fiables [PM18]. Judea Pearl l'exprime précisément en ces termes : « *It is critical to realize that data are profoundly dumb about causal relationships* ») [PM18]. Cette citation explique littéralement que les données étant « *profondément stupides* », elles ne possèdent pas intrinsèquement la capacité de saisir ou d'inférer les mécanismes causaux sous-jacents (cause et effet) de manière autonome. Dans ce contexte, il est impératif de développer des méthodes d'apprentissage automatique robustes et flexibles pour répondre aux exigences posées par des données dites « complexes » — qu'elles soient parcimonieuses (sparses), de très grande dimension (« *curse of dimensionality* »), difficiles à interpréter, coûteuses en calcul, ou dans le cadre d'environnement non stationnaire.

Ce travail d'Habilitation à Diriger des Recherches (HDR) se concentre précisément sur l'application de l'apprentissage automatique à la génération et l'exploration pour données complexes, avec un accent particulier sur les problématiques d'adaptation et de contextualisation. Les approches explorées dans ce mémoire s'appuient sur des méthodes de pointe issues de l'**apprentissage par renforcement**, de l'**apprentissage profond**, et des **méthodes évolutionnaires**. Ces techniques sont appliquées à divers domaines transversaux, tels que les systèmes de recommandation, la classification dans le domaine médical, l'estimation de paramètres dans des systèmes dynamiques, et les tâches de génération. Cette HDR s'appuie à la fois sur des implémentations, des idées et des travaux personnels, ainsi que sur des contributions d'équipe, que ce soit dans le cadre de collaborations institutionnelles ou de co-encadrements de thèse. Ainsi, j'utiliserai par la suite le « nous » pour désigner les travaux exposés dans cette HDR.

L'optimisation multicritère appliquée à l'apprentissage (voir Partie I) constitue un aspect important de mes recherches, en particulier l'adaptation des méthodes d'apprentissage par renforcement aux environnements non stationnaires et aux données parcimonieuses dans un cadre multi-objectif (voir Chapitre 1). Dans le domaine de l'**apprentissage par renforcement**, Richard Sutton et Andrew Barto soulignent que la capacité d'un agent à apprendre des actions optimales à partir de ses interactions avec l'environnement est la clé de l'IA [SB98]. C'est dans cette optique que nous avons particulièrement travaillé sur l'optimisation multicritère appliquée aux systèmes de recommandation, en nous appuyant sur des algorithmes de bandits-manchots. Nous avons notamment proposé de nouvelles méthodes de type portfolio d'algorithmes en s'appuyant sur les algorithmes de bandits-manchots contextuels et combinatoires. Ces méthodes visaient à améliorer à la fois la précision et la diversité des recommandations, tout en s'adaptant à des environnements non stationnaires, c.-à-d., en constante évolution. En complément de cet axe, j'ai travaillé sur le problème de sélection multi-objectif d'attributs, notamment dans des applications médicales. Pour cela, j'ai employé des **méthodes évolutionnaires** (voir Chapitre 2), qui jouent un rôle essentiel dans ce domaine. David E. Goldberg, un expert en algorithmes génétiques, explique dès 1989 que les méthodes évolutionnaires fournissent une approche puissante et flexible pour explorer des espaces de solution complexes et multidimensionnels [Gol89]. Nous avons ainsi mis en évidence l'utilisation de ces méthodes pour améliorer la sélection d'attributs en médecine, combinant algorithmes génétiques et méthodes de calcul de fitness multicritère pour optimiser les décisions médicales.

En parallèle de ces recherches centrées sur des problématiques d'optimisation multicritère, je me suis penché sur l'**apprentissage profond** (voir Partie II), qui représente un volet important et incontournable de la recherche en IA. Yann LeCun, l'un des pionniers du domaine, a souligné que cette approche a révolutionné notre capacité à extraire automatiquement des caractéristiques complexes et à modéliser des relations non linéaires à partir de données non structurées [LBH15]. En 2024, Geoffrey Hinton a même été co-récepteur du prix Nobel de physique pour ses travaux fondamentaux sur les réseaux de neurones artificiels et l'apprentissage automatique, soulignant ainsi l'importance de cette discipline dans les avancées technologiques majeures, telles que la reconnaissance d'images, les assistants vocaux et l'IA générative. Pour cause, l'apprentissage profond s'est avéré particulièrement puissant dans de nombreux travaux de la dernière décennie, p.ex., les travaux de Goodfellow et al. [Goo16] ont démontré comment les réseaux de neurones profonds peuvent capturer des représentations hiérarchiques de données complexes, permettant des avancées significatives dans la reconnaissance d'images et le traitement du langage naturel. De même, Radford et al. [Rad+19] ont illustré l'efficacité des modèles de type *GPT* pour la génération de texte et de musique, en exploitant la capacité de ces modèles à comprendre des structures séquentielles et contextuelles de grande envergure. Inspirés par ces résultats et les promesses qu'ils représentent, nous avons employé l'apprentissage profond dans divers travaux : 1) Pour l'estimation de paramètres dans les systèmes dynamiques (voir Chapitre 3) via l'usage de *MultiLayer Perceptron (MLP)*, de *Gated recurrent units (GRU)*, de *Convolutional Neural Network CNN (CNN)* ou encore de *CNN-Transformers* ; 2) Pour la génération de musique symbolique (voir Chapitre 4) à

l'aide de modèles reposant sur les Transformers (*GPT2* ou *BERT*). Nos travaux ont de nouveau démontré la pertinence des réseaux de neurones profonds. Cependant, comme tout chercheur en apprentissage automatique, nous nous sommes confrontés à des problématiques liées aux biais ou à la qualité des données. Cela a pu inclure, par exemple, des problèmes de représentation structurée des données d'apprentissage, de parcimonie (sparsity) des données, de déséquilibre dans l'ensemble d'entraînement, de dérive conceptuelle (concept drift) causée par la non-stationnarité, ou encore de changement de concept (concept shift) résultant, par exemple, d'une perte d'informations provenant de capteurs. Yann Lecun affirme à ce titre que des algorithmes entraînés sur des données biaisées seraient, eux aussi, biaisés bien que les modèles sur lesquels ils reposent ne l'étaient pas (« *Algorithms trained on biased data are biased* »). C'est pourquoi, pour optimiser les performances de ces modèles, nos contributions ne se sont pas limitées qu'à leur simple utilisation ou amélioration. Nous avons également concentré nos efforts sur la structuration de l'information fournie aux réseaux de neurones ou aux modèles génératifs, notamment à travers l'amélioration de processus clés tels que la tokenisation de la musique symbolique (Chapitre 4). De plus, nous avons intégré des informations complémentaires « complexes » (Chapitre 3), comme pour les réseaux informés par la physique (*Physics-Informed*), afin d'améliorer à la fois la précision et l'efficacité de ces modèles.

Pour aborder ces différents domaines et ces diverses contributions, mon manuscrit s'articule autour de deux grandes parties. La première partie aborde l'adaptation et la contextualisation dans le cadre de l'optimisation multicritère appliquée à l'apprentissage, tandis que la seconde traite de l'approximation et de la représentation de données « complexes », dans le cadre de tâches d'estimation ou de génération, en utilisant des méthodes d'apprentissage profond. Enfin, après avoir dressé un bilan de mes travaux de ces cinq dernières années, je présente mes perspectives de recherche à travers le projet ^G-*GENOCOD* (*Graph-Generation for Novel Compound Discovery*), qui vise à développer de nouvelles approches d'IA pour la génération de composés moléculaires. L'ensemble de ces travaux contribue à faire progresser l'apprentissage automatique tout en ouvrant de nouvelles pistes pour la recherche interdisciplinaire.

Organisation du mémoire

Liens vers les sites web

Ce mémoire fait référence à un certain nombre de sites web.

Les liens de ces sites sont donnés « *directs* » dans le corps du mémoire, c'est-à-dire selon leur accessibilité dans les années 2024, 2025. En revanche, afin de pérenniser leur consultation, chacun de ces liens est répertorié, « *archivé* » dans l'Annexe C. Ainsi, les liens se rapportant aux différents sites sont disponibles dans le Tableau C.1.

Organisation du manuscrit

Avant d'aborder le détail de mes contributions, ce manuscrit propose une présentation synthétique de mon profil en préambule. Il met en lumière :

1. Le résumé de mon profil ;
2. Mon Curriculum Vitæ ;
3. Mes activités d'enseignement et mes charges administratives associées
4. Mes activités de recherche en expliquant dans le détail l'articulation logique et chronologique de mes implications et contributions ;
5. Mes responsabilités scientifiques ;
6. Mes encadrements doctoraux et post-doctoraux ;
7. Ma production scientifique rappelant l'ensemble de mes travaux publiés ;
8. Une réflexion prospective synthétique sur mes perspectives de recherche et de carrière.

Après avoir rappelé mon profil, le reste du mémoire est dédié à la description de mes travaux et de mes contributions.

Les travaux présentés se concentrent sur l'apprentissage automatique pour la génération et l'exploration. Ils traitent plus particulièrement des problématiques d'adaptation et de contextualisation, au regard de données d'entraînement ou à prédire, dites « complexes ».

Ces contributions sont détaillées en deux parties dans le manuscrit :

1. Partie I : **Optimisation multicritère appliquée à l'apprentissage**

La première partie des travaux aborde les problématiques d'adaptation et de contextualisation dans le cadre de l'optimisation multicritère appliquée à l'apprentissage. Elle se compose de deux chapitres :

(a) Chapitre 1 : **Apprentissage par renforcement et systèmes de recommandation multicritères.**

Ce chapitre se concentre sur l'apprentissage par renforcement, en particulier les problèmes d'optimisation multicritère dans les systèmes de recommandation. Il présente plus précisément une suite de nouvelles méthodes multicritères de type portfolio d'algorithmes développée pour améliorer la précision, la diversité des recommandations et pour s'adapter à des environnements non stationnaires ou à des données parcimonieuses. Ces méthodes emploient des algorithmes de bandits-manchots contextuels, non contextuels, et combinatoires. Dans ce volet nous évoquerons également de manière complémentaire (voir Annexe A.1), l'influence des fonctions de récompense sur les performances des algorithmes de bandits-manchots. Ces travaux proposent de nouvelles méthodes de sélection des retours utilisateur et de calcul de récompense plus performantes permettant de mieux s'adapter à des retours souvent partiels.

(b) Chapitre 2 : **Optimisation multicritère pour la sélection d'attributs médicaux.**

Ce chapitre est consacré aux méthodes évolutionnaires et à leur application à la sélection multi-objectif d'attributs dans le domaine médical. Il présente notre méthode proposant un algorithme génétique combiné à une méthode de calcul de fitness multicritère, basée sur un *Compass*, afin de résoudre des problèmes complexes de sélection d'attributs médicaux. Cette approche, qui repose sur les méthodes présentées dans le premier chapitre, a permis d'améliorer les processus de décision médicale tout en surpassant des méthodes traditionnelles concurrentes.

2. Partie II : **Apprentissage profond et données complexes pour l'estimation et la génération**

La seconde partie des travaux aborde les problématiques d'estimation et de représentation de données d'entraînement complexes dans le cadre de l'apprentissage profond, allant des méthodes génératives aux méthodes d'estimation. Elle se compose de deux chapitres :

— Chapitre 3 : **Évaluation de comportements dynamiques non linéaires : application aux véhicules terrestres.**

Ce chapitre explore l'application de l'apprentissage profond et de méthodes bayésiennes pour l'estimation de paramètres dans les systèmes dynamiques, en particulier dans le cadre des dynamiques de véhicules terrestres. Il propose l'utilisation de méthodes d'inférence bayésienne et de méthodes hybrides (*Data-based* et *Model-based*) incluant la mise en œuvre de réseaux de neurones profonds et de méthodes d'apprentissage par transfert (*Transfert Learning*) pour améliorer la précision des estimations de paramètres. De plus, ce chapitre propose une approche d'apprentissage profond visant à apprendre une solution d'optimisation *MHE* (*Moving Horizon Estimation* : Estimation à Horizon Glissant) hors ligne, permettant au modèle de fournir des estimations plus rapides et plus précises à partir des observations du système, sans nécessiter d'optimisation en ligne.

— Chapitre 4 : **Optimisation des représentations symboliques musicales pour l'apprentissage profond**

Ce chapitre se concentre sur l'apprentissage profond reposant sur les modèles à base de *Transformers* (modèles *GPT2* ou *BERT*) pour la génération de musique symbolique. L'accent est mis sur l'amélioration du processus de *tokenisation* de la musique, crucial pour la création de nouvelles compositions musicales grâce aux modèles génératifs. Pour ce faire, ce chapitre présente d'abord une contribution logicielle, *MidiTok*, destinée à uniformiser la tokenisation des fichiers MIDI dans le cadre de la musique symbolique. Ensuite, il examine l'impact du temps et de la durée des notes sur le processus de tokenisation, des aspects critiques puisque les représentations choisies pour ces dimensions peuvent affecter à la fois la taille des séquences de tokens, la complexité des modèles et la qualité des résultats en génération ou en classification musicale. Enfin, ce chapitre propose une méthode d'amélioration de la tokenisation via l'utilisation du *Byte Pair Encoding* (*BPE*), permettant de réduire la

longueur des séquences tout en augmentant la taille du vocabulaire, améliorant ainsi la qualité des générations.

Conclusion et perspectives de recherche. Enfin, après avoir rappelé mon bilan à date, je décris mes perspectives de recherches au sein d'une troisième partie, au travers du projet *G-GENOCOD* (*Graph-Generation for Novel Compound Discovery*), lancé en 2024. Ce projet, soutenu par la région Pays de la Loire via le dispositif « Étoiles Montantes », vise à développer de nouvelles méthodes d'IA pour la génération moléculaire, afin de découvrir de nouveaux composés. Pour cela, nous explorons actuellement soit une combinaison de méthodes d'apprentissage par renforcement et de méthodes évolutionnaires, soit l'utilisation de l'apprentissage profond basé sur les *Transformers*. Ce projet capitalise sur la convergence des compétences scientifiques et des méthodes issues de mes projets précédents, tout en élargissant mes travaux actuels et ouvrant de nouvelles perspectives de recherche et d'application.

Annexes. Ce mémoire comporte quatre annexes permettant de mieux préciser ou détailler certains chapitres.

- l'Annexe A a une utilité double : 1) Elle décrit nos contributions complémentaires ou connexes au sujet de cette HDR ; 2) Elle donne quelques détails supplémentaires sur certaines contributions déjà décrites dans ce mémoire ;
- l'Annexe B permet de consulter différents pseudo-codes d'algorithmes de l'état de l'art présentés dans le Chapitre 1 ;
- l'Annexe C répertorie les liens archivés de l'ensemble des sites web cités dans ce mémoire.

PRÉSENTATION SYNTHÉTIQUE

Résumé du profil

Diplômé en 2006 d'un Master à finalité professionnelle en Génie Physiologique-Informatique (*GPHY*) de l'Université de Poitiers, mon parcours professionnel débute en tant qu'ingénieur au sein d'entreprises de services du numérique (ESN) à Paris, où j'ai exercé de 2006 à 2010.

À partir de septembre 2010, j'ai occupé des fonctions de chargé de cours à temps partiel (vacataire) dans quatre universités distinctes : Angers, Poitiers, La Rochelle et Rouen. Un an plus tard, en 2011, j'ai été recruté par l'ESAIP en tant qu'enseignant en informatique et responsable de la filière ingénieur Informatique et Réseaux. J'ai occupé cette fonction jusqu'en juin 2016 date à laquelle j'ai rejoint l'ESEO pour un poste d'enseignant en informatique (210 HETD/an). Durant cette période, j'ai effectué en parallèle un travail de thèse de 3 ans en tant que doctorant en informatique au LERIA (Université d'Angers) sous la direction de Tassadit Amghar et d'Olivier Camp. Ma thèse, soutenue en 2019, abordait des thématiques portant sur l'apprentissage par renforcement et sur les systèmes de recommandation. J'ai ensuite prolongé mon engagement en tant qu'enseignant-chercheur pendant près d'un an (2019-2020) au sein de l'ESEO avant de rejoindre pleinement l'Université d'Angers à l'IUT en septembre 2020.

Depuis 2020, je suis donc Maître de conférences en Informatique à l'IUT d'Angers-Cholet. J'y assure mes enseignements dans le Département GEII (Génie Électrique et Informatique Industrielle) principalement dans les domaines des bases de données, de la programmation, des réseaux informatiques et industriels, et de la cybersécurité. J'y ai également eu des implications administratives en tant que Directeur du parcours Électronique & Systèmes Embarqués. Concernant ma partie recherche, je suis rattaché au LERIA où j'y réalise mes travaux de section 27 majoritairement au sein de l'axe ARC : Apprentissage Artificiel et Représentation des Connaissances.

Curriculum Vitæ

Identité

Nom :	GUTOWSKI
Prénom :	Nicolas
Date et lieu de naissance :	9 novembre 1982, à Senlis (Oise)
Nationalité :	Française
Situation de famille :	Marié, 2 enfants
Dernier diplôme obtenu :	Doctorat en informatique de l'Université d'Angers
Profession :	Maître de conférences en Informatique
Coordonnées professionnelles (composante) :	IUT d'Angers-Cholet 4 Boulevard de Lavoisier 49000 ANGERS
Coordonnées professionnelles (laboratoire) :	UFR Sciences, LERIA 2 Boulevard de Lavoisier 49000 ANGERS
Courriel :	nicolas.gutowski@univ-angers.fr
Page personnelle :	https://ngutowski.fr/

Formation et titres universitaires

2016-2019 : Doctorat en Informatique de l'Université d'Angers.

Titre : « *Recommandation contextuelle de services : Application à la recommandation d'événements culturels dans la ville intelligente.* »

Jury :

<i>Président du Jury :</i>	Pr. Éric Monfroy (Université d'Angers)
<i>Directeur :</i>	Dr. Tassadit Amghar (Université d'Angers)
<i>Co-encadrant :</i>	Dr. Olivier Camp (ESEO)
<i>Rapporteurs :</i>	Pr. Allel HadjAli (ISAE-ENSMA) Dr. Armelle Brun (Université de Lorraine)
<i>Examineur :</i>	Pr. Florence Sèdes (Université Paul Sabatier) Pr. Bruno Zanuttini (Université de Caen) Pr. Éric Monfroy (Université d'Angers) Dr. Raphaël Féraud (Orange Labs)

2005-2006 : Master à finalité professionnelle en Génie Physiologique-Informatique,
de l'Université de Poitiers - mention Assez-Bien

2004-2005 : Maîtrise en Sciences et Technologies de l'IUP Génie Physiologique-Informatique,
de l'Université de Poitiers - mention Assez-Bien

2003-2004 : Licence IUP Génie Physiologique-Informatique, de l'Université de Poitiers - mention Bien

2002-2003 : DEUG IUP Génie Physiologique-Informatique de l'Université de Poitiers - mention Assez-Bien

2001-2002 : Classes Préparatoires aux Grandes Écoles, filière *BCPST : Biologie, Chimie, Physique, Sciences de la Terre* au Lycée Camille Guérin à Poitiers

2000-2001 : Baccalauréat Scientifique, au Lycée Merleau Ponty à Rochefort - mention Bien.

Expérience professionnelle

Depuis septembre 2020 : Maître de conférences en Informatique à l'**IUT** d'Angers-Cholet (Département GEII, Laboratoire LERIA).

J'enseigne à l'IUT avec une charge annuelle moyenne comprise entre 192 et 300 HETD (hors première année de stage fixée à 160 HETD), et je mène mes recherches au LERIA dans l'axe ARC (Apprentissage Artificiel et Représentation des Connaissances). J'ai été responsable du parcours Électronique et Système Embarqué (ÉSE) du BUT GEII d'avril 2022 à août 2024.

Novembre 2019 - Août 2020 : Enseignant-Chercheur à l'**ESEO**, Angers (<https://www.eseo.fr>).

L'ESEO est une école d'ingénieur généraliste en cinq ans habilitée par la Commission des Titres d'Ingénieur (CTI) depuis 1962. J'y ai enseigné pour une charge annuelle d'enseignement en informatique de 210 HETD et j'y ai effectué ma recherche sur les systèmes de recommandation contextuels reposant sur les bandits-manchots. Cette recherche était le prolongement de mes travaux de thèse. J'étais également le responsable de l'axe de recherche en « Ville Intelligente » de l'école.

Juin 2016 - Novembre 2019 : Enseignant à l'**ESEO**, Angers.

J'ai enseigné à l'ESEO pour une charge annuelle de 210 HETD aussi bien dans les enseignements de tronc commun du cycle ingénieur qu'en option de dernière année. Je suis également intervenu en troisième année du cursus Bachelor.

Durant ces trois années en tant qu'enseignant à l'ESEO, j'ai eu l'opportunité de réaliser mon doctorat ayant pour sujet de recherche la recommandation contextuelle de services.

Septembre 2011 - Juin 2016 : Responsable du cycle ingénieur Informatique et Réseaux à l'**ESAIP**, Angers (<https://www.esaip.org/>).

L'ESAIP est une école d'ingénieur en Informatique en cinq ans habilitée par la Commission des Titres d'Ingénieur (CTI) depuis 1998. Durant ces 5 années passées à l'ESAIP, j'ai assuré la responsabilité du cycle ingénieur Informatique et Réseau qui englobait principalement les missions de suivi des étudiants, de management du département, de développement et d'accréditation CTI des programmes de la filière Informatique et

Réseaux. J'y ai également assuré une charge d'enseignements allant de 215 à 270 HTD par an en France mais aussi en Europe (en langue anglaise) dans le cadre de mobilités ERASMUS pour les universités partenaires de l'ESAIP.

Avril 2010 - Septembre 2011 : Formateur indépendant, Vacataire.

Durant cette période de transition vers le monde de l'enseignement supérieur, j'ai pu assurer des missions en tant qu'enseignant vacataire en M2 Génie Physiologique-Informatique à l'université de Poitiers, en M2 Bioinformatique à l'université de Rouen, en M1 à l'ISSBA (Polytech) Angers, en Licence Professionnelle GEB (Gestion Électrique du Bâtiment) à l'université d'Angers et en L1 (Droit et LEA) à l'université de La Rochelle. J'ai également pu effectuer des missions de formation professionnelle pour les entreprises notamment pour le compte de la CCI. Ces expériences en vacation dans différentes universités m'ont permis d'acquérir une grande adaptabilité dans l'élaboration de cours, TD et TP de domaines très variés et dans un temps restreint.

Août 2006 - Avril 2010 : Consultant en informatique en ESN (SSII) à Altran CIS puis Héraclès, Paris.

Durant ces années j'ai pu assurer diverses missions de conseil dans le domaine des TICs (Technologies de l'Information et de la Communication) mais aussi des SI (Systèmes d'Information). Entre autres j'ai pu assurer des missions de Data analyst (avec étude de la mise en place d'indicateurs de performance) ou d'homologateur dans des entreprises du domaines de la banque, de la finance et de l'assurance. Ces expériences m'ont donné la culture du service et ont forgé ma capacité à répondre rapidement et avec efficacité aux besoins des structures pour lesquelles je travaille.

Activités d'enseignement et charges administratives associées

Activités d'enseignement

L'intérêt que je porte à l'enseignement et à la formation, m'a conduit à intervenir dès 2010 en tant que vacataire dans diverses formations de l'enseignement supérieur (Université de Poitiers, Université de La Rochelle, Université de Rouen, Université d'Angers).

Dans le cadre de mes activités en tant qu'enseignant à temps plein, en complément de mon volume horaire annuel, j'ai également été amené à réaliser des mobilités d'enseignements ERASMUS en anglais (Aalborg au Danemark et Darmstadt en Allemagne en 2013, Madrid et Valladolid en Espagne en 2014 et 2015). Cela m'a permis de développer des partenariats avec des universités européennes mais aussi d'échanger sur nos bonnes pratiques pédagogiques et scientifiques.

Ainsi depuis 2010, j'ai dispensé plus de 3000 heures d'enseignements dans des cursus variés (voir Tableau 1). Le Tableau 1 témoigne de mon adaptabilité à enseigner dans différentes filières de différents niveaux allant de Bac+1 à 5.

Depuis 2020, j'assure des enseignements dans ma composante IUT essentiellement et plus particulièrement dans le département GEII : Génie Électrique et Informatique Industrielle

(c.-à-d., BUT GEII et Licence Professionnelle Réseaux & Télécoms, licence suspendue puis fermée en 09/2024).

Formations techniques (Bachelor, DUT, BTS, LP)	
DUT R&T	DUT Réseaux & Télécoms (IUT de Chatellerauld)
TSRIT	Technicien Supérieur en Réseaux Informatiques et Télécommunications (ESAIP)
BTS SN / SIO	BTS Systèmes Numériques / BTS Services Informatique aux Organisations
B3 ESEO	Bachelor Numérique (ESEO)
BUT GEII	BUT Génie Électrique et Informatique Industrielle (IUT d'Angers-Cholet)
LP RT	Licence Professionnelle Réseaux et Télécoms (IUT d'Angers-Cholet)
LP MSA	Licence Professionnelle Maintenance des Systèmes Automatisés (IUT d'Angers-Cholet)
LP-GEB	Licence Professionnelle Gestion Électrique du Bâtiment (Université d'Angers)
LP-ECV	Licence Professionnelle Essai Clinique & Validation (Université de Poitiers)
Formations Licences/Masters	
L1	Licence (Droit ou LEA) (Université de La Rochelle)
M1-STIS	Master Sciences, Technologies, et Ingénierie de la Santé (ISSBA)
CPI2, CPI3	Master Chef de Projet Informatique (ESAIP)
M1/M2-GPHY	Master Génie Physiologique-Informatique (Université de Poitiers)
M2-BIO	Master Bioinformatique par alternance (Université de Rouen)
M2-IIA	Master Manager de l'Ingénierie Informatique (CCI de Laval)
M2-MIASHS	Master Mathématiques et Informatique Appliquées aux SHS (UCO Angers)
Formations ingénieurs habilitées par la CTI	
IR1, IR2, IR3	1 ^{ère} , 2 ^{ème} et 3 ^{ème} années Ingénieur Informatique & Réseaux (ESAIP)
I1, I2, I3	1 ^{ère} , 2 ^{ème} et 3 ^{ème} années Ingénieur ESEO
P2 ESAIP	2 ^{ème} année de classe préparatoire de l'ESAIP
ITII3	3 ^{ème} année Ingénieur Informatique Industrielle ESEO (avec l'ITII)
A3 ESEO	3 ^{ème} année Ingénieur ESEO (avec la CCI)
SEP3	3 ^{ème} année Ingénieur Prévention des risques et Sécurité (ESAIP)
I3-IRSI	3 ^{ème} année Ingénieur Intégration des Réseaux et des SI (EIGSI)
I3-QA	3 ^{ème} année Ingénieur Qualité, Innovation, Fiabilité (ISTIA/Polytech)
Formations à l'internationale	
Computer Science (M.Sc.)	Hochschule Darmstadt (h_da) - Université de Darmstadt (Allemagne)
Computer Science (B.Sc.)	University College of Northern (UCN, Aalborg, Danemark)
Engineering (B.Sc.)	Université de P. Comillas (Madrid, Espagne)
Computer Engineering (M.Sc.)	Université de Valladolid (UVa, Espagne)

TABLE 1 – Formations dans lesquelles je suis intervenu depuis 2010

Ci-dessous je présente une synthèse des principaux enseignements dispensés par période :

- **Depuis 2020** : Enseignement à l'IUT d'Angers-Cholet, avec un volume total annuel allant de 192 à 300 heures équivalent TD. Les disciplines enseignées concernent principalement les réseaux, la cybersécurité, les bases de données, la gestion de projets, le Green IT. Sont inclus dans ce volume, mes suivis de stage et d'alternance, ma responsabilité de parcours (36h).
- **2016-2020** : Enseignement à l'ESEO, avec un volume annuel d'environ 210 heures équivalent TD. Les disciplines enseignées concernent principalement le développement JAVA, l'UML, la programmation orientée objet, les bases de données, le développement Web, l'informatique décisionnelle, la gestion de projets.
- **2011-2016** : Enseignement à l'ESAIP, avec un volume annuel d'environ 270 heures équivalent TD. Les disciplines enseignées concernent principalement le Green IT, les bases de données, le développement Web, l'informatique décisionnelle, la gestion de projets. Durant la période à l'ESAIP j'ai également enseigné en teaching mobility ERASMUS en Green IT (Informatique durable) et développement Web.
- **2010-2011** : Enseignement en tant que vacataire dans différentes universités avec un volume annuel correspondant à un service de Maître de conférences. Les disciplines enseignées concernent principalement le développement Web, les bases de données, la gestion de projets.

Responsabilité administratives :

- **03/2022 - 08/2024** : Responsable du parcours Électronique et Systèmes Embarqués du Département GEII (IUT d'Angers-Cholet). Cette mission correspond principalement à la gestion des alternances/stages des étudiants, aux relations avec les entreprises, à la gestion d'intervenants extérieurs, à la validation des contenus de semestres à l'étranger, aux préconisations d'évolution du programme de formation.
- **09/2011 - 06/2016** : Responsable du cycle ingénieur en Informatique et Réseaux à l'ESAIP. Cette mission correspond principalement à celle d'un chef de département ou responsable d'un cycle complet de formation. Il inclut les missions de responsable des programmes (parcours) et de direction des études des années bac+3 à 5 mais aussi de recrutement des enseignants permanents ou des vacataires.

Activités de recherche

Vocation en recherche - J'ai effectué mes premiers pas dans la recherche scientifique lors de stages effectués en 2003 et 2004, dans le cadre de mon DEUG puis de ma Licence. Le premier stage a été réalisé à l'INSERM à Marseille, et le suivant au centre de recherche Nemours à Wilmington (Delaware), aux États-Unis. À cette époque, j'étais inscrit dans un cursus orienté biologie et informatique à l'IUP de Poitiers (Institut Universitaire Professionnalisé). Cet institut

délivrait in fine un Master professionnel, que j'ai obtenu en 2006. Bien que déjà très attiré par la recherche, les circonstances et aspirations de l'époque ne m'ont pas conduit à poursuivre directement avec un doctorat. Toutefois, l'idée d'entreprendre une thèse resta présente et s'imposa progressivement à moi comme une évidence.

Après dix années d'expérience professionnelle, dont six ans dédiés à l'enseignement supérieur, j'ai ainsi décidé en 2016 de reprendre le chemin de la recherche académique en commençant une thèse au LERIA de l'Université d'Angers, au sein de l'équipe Apprentissage et Représentation des Connaissances. Ma thèse portait sur la recommandation contextuelle de services, avec une application spécifique aux événements culturels dans les villes intelligentes, utilisant principalement les bandits-manchots contextuels.

Après l'obtention de mon doctorat en novembre 2019, j'ai diversifié mes activités de recherche principalement autour de trois axes : 1) les systèmes de recommandation basés sur les bandits manchots ; 2) l'optimisation multicritère ; 3) l'apprentissage profond (Deep Learning). Ces axes bien que présentés séparément par la suite, se croisent souvent dans mes projets et s'enrichissent mutuellement dans le but d'améliorer les performances.

Rappel des travaux de thèse (2016-2019)

Ma thèse avait porté sur les algorithmes de bandits-manchots pour les systèmes de recommandation sensibles au contexte [Gut19]. Afin de répondre aux enjeux de cette thématique, les travaux et contributions s'étaient organisés autour de trois axes :

1. Les systèmes de recommandation ;
2. Les algorithmes de bandits-manchots contextuels et non contextuels (Contextual Multi-Armed Bandit : CMAB, et Multi-Armed Bandit : MAB) ;
3. Le contexte.

Les travaux des deux premières années de thèse avaient permis de candidater et d'aboutir à l'obtention d'une subvention régionale dans le cadre d'un appel d'offre RFI (Recherche Formation Innovation du dispositif Atlanstic 2020). Ce projet de recherche appliquée, Event-AI, avait pour objectif le développement d'une application mobile (scéno) pour la visualisation et la recommandation d'événements culturels à des utilisateurs dans la ville d'Angers. Le projet Event-AI débuté en février 2019, avait permis d'intégrer à l'application certaines méthodes évaluées de manières empiriques durant la thèse, et ainsi de les évaluer en ligne. En début de thèse (2016-2017), une première contribution en journal [Gut+17], étendue en chapitre de livre en 2018 [Sed18] avait porté sur les systèmes de recommandation de services dans la ville intelligente. Cette contribution avait porté sur la capture et le raisonnement contextuel pour les systèmes de recommandation, à partir de connexions d'utilisateurs mobiles au Wi-Fi urbain. Cette étude avait été ensuite étendue en 2018 sur des travaux en géopartitionnement afin d'inférer du contexte géographique et de l'intégrer dans le système de recommandation d'événements culturels du projet Event-AI. Cette contribution avait donné lieu à une publication en conférence internationale [Gut+19c]. Enfin, les principales contributions avaient porté sur les algorithmes de CMAB pour la recommandation, à savoir trois points :

- la diversification des recommandations [Gut+19a] : Soit via la modification d'un algorithme de CMAB : LinUCB [Li+10]; Soit via une approche porte-folio d'algorithmes de MAB et CMAB [Gut+19b].
- la précision individuelle des recommandations effectuées par les algorithmes de MAB et CMAB [Gut+19a; Gut+19b];
- l'enrichissement dynamique de contexte des algorithmes de CMAB [Gut+18; Gut+19d] : notamment LinUCB (Linear Upper Confidence Bound) [Li+10] et Contextual Thompson Sampling [AG13b].

Rappel des thématiques de recherche du LERIA

Trois thématiques de recherche sont portées par le LERIA :

1. Apprentissage et Représentation des Connaissances (ARC);
2. Métaheuristiques et Optimisation Combinatoire (MOC);
3. Raisonnement dans l'Incertain et Contraintes (RIC).

Lors de mon intégration au laboratoire en 2020, j'ai été rattaché à la thématique ARC. En revanche, certains de mes travaux se situent à l'intersection des thématiques ARC et MOC.

Les travaux du thème ARC développent des solutions basées sur des modèles de représentation des connaissances et des méthodes d'apprentissage artificiel pour analyser et comprendre des problèmes scientifiques ou applicatifs. En matière d'apprentissage, certaines de ces recherches portent sur les systèmes de recommandation et l'inférence de connaissances en biologie et chimie moléculaire. En collaboration avec le thème MOC, ils intègrent également des méthodes d'apprentissage et de fouille de données dans les recherches méta-heuristiques. Au sein de cette thématique par exemple : les systèmes de recommandation utilisent des algorithmes de bandits-manchots pour la recommandation contextuelle et l'apprentissage multicritère; en chimie moléculaire, des méthodes d'IA sont appliquées pour la découverte de nouvelles molécules, avec un générateur moléculaire basé sur un algorithme évolutionnaire.

Depuis 2020, mon intégration au sein du LERIA s'est effectuée progressivement sur la thématique ARC. De ce fait, mes travaux ont tout au long porté sur l'apprentissage par renforcement puis sur l'apprentissage automatique et profond, combinés aux domaines de recherche de la thématique MOC, notamment la résolution de problèmes d'optimisation et l'utilisation de méthodes évolutionnaires.

Mes axes de recherche : articulation logique et chronologique

À la suite de ma thèse soutenue le 4 novembre 2019, j'ai poursuivi et consolidé mes compétences dans le domaine de l'apprentissage par renforcement, en me concentrant sur des problèmes d'optimisation bi-objectif inhérents aux systèmes de recommandation (de fin 2019 à 2024). Parallèlement, j'ai développé des compétences en méthodes génératives dans le domaine de l'apprentissage profond pour la génération de musique symbolique (de 2021 à 2024).

J'ai également renforcé mes compétences en apprentissage profond pour l'estimation de paramètres dans les systèmes dynamiques (de 2022 à 2025). En outre, j'ai travaillé dans le domaine de recherche de l'algorithmique évolutionnaire, notamment les algorithmes génétiques, et plus généralement de l'optimisation, en me concentrant sur la sélection multicritère d'attributs dans le domaine médical (2021-2022).

Enfin, à la convergence de l'ensemble de ces domaines, j'ai entamé des travaux appliqués à la génération moléculaire (de 2024 à 2027). Ces travaux abordent le problème sous deux angles différents : la génération en utilisant des méthodes issues de l'apprentissage profond, notamment les *Transformers*, et des méthodes hybrides combinant des techniques évolutionnaires et d'apprentissage par renforcement (Bandits-Manchots, *Q-Learning*).

Afin d'illustrer le positionnement de mes travaux à travers ces différents domaines, je propose dans un premier temps une visualisation présentant l'articulation et les intersections entre ces différents domaines de recherches et les domaines applicatifs sur lesquels j'ai travaillé. La Figure 1 présente un diagramme de Venn illustrant les intersections entre les quatre domaines de recherche sur lesquels j'ai travaillé, les applications, et les tâches.

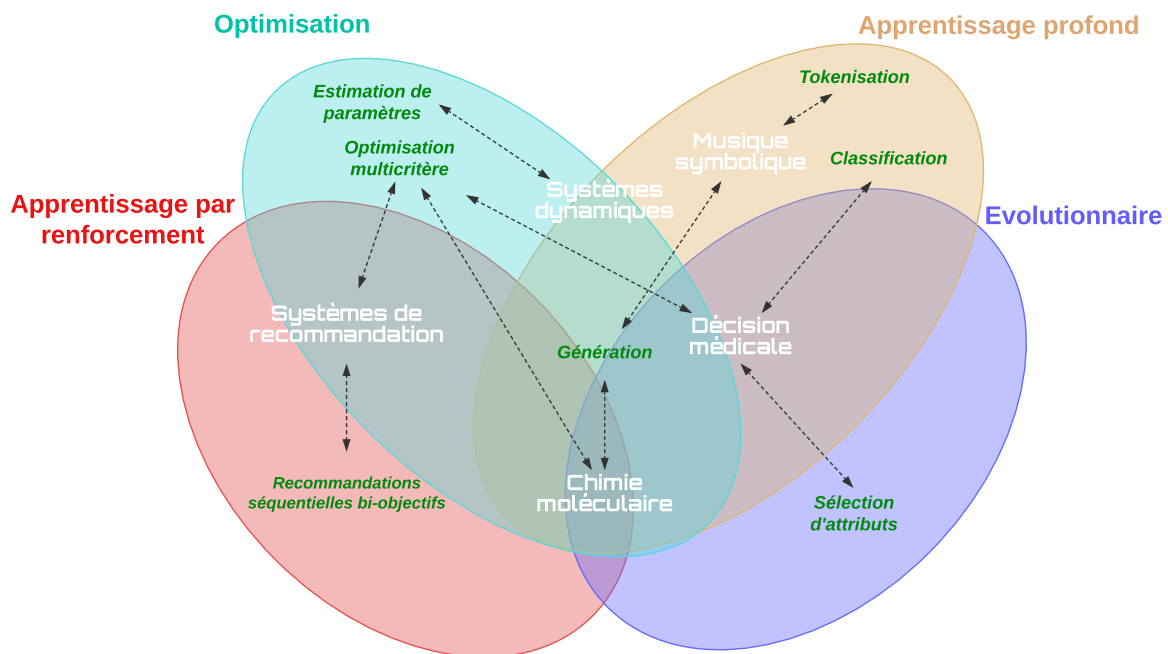


FIGURE 1 – Articulation logique des thématiques appliquées selon les domaines de recherche et les tâches effectuées

Plus précisément, ce diagramme de Venn permet de visualiser :

1. l'ensemble des quatre domaines de recherche (apprentissage par renforcement, optimisation, apprentissage profond, et domaine de l'évolutionnaire) selon 4 couleurs distinctes
2. les domaines applicatifs en blanc ;
3. les principales tâches sur lesquelles je me suis plus particulièrement focalisé en vert.

Le diagramme met en évidence les relations et la synergie entre les différents domaines de recherche, illustrant comment ils s'entrecroisent dans divers projets et applications, notamment dans la génération moléculaire, où se concentrent la majorité des compétences et méthodes issues de chaque domaine.

Ainsi, on remarquera les intersections suivantes :

- Systèmes de recommandation : Convergence entre apprentissage par renforcement et optimisation via le processus de recommandations séquentielles bi-objectifs ;
- Systèmes dynamiques : Intersection entre optimisation et apprentissage profond, en se focalisant sur l'estimation de paramètres ;
- Décision médicale : Intersection entre apprentissage profond, et les méthodes évolutionnaires, avec un focus sur la sélection d'attributs mutli-objectif ;
- Musique symbolique : Tokenisation et génération, relevant de l'apprentissage automatique et profond ;
- Chimie moléculaire : Représente la convergence des quatre domaines, avec en perspective des travaux de génération utilisant d'une part des méthodes de l'apprentissage profond et d'autres parts des méthodes combinées d'apprentissage par renforcement et évolutionnaires.

Pour mieux situer mes activités dans une dimension temporelle, la Figure 2 illustre l'enchaînement et l'évolution de mes travaux de recherche à travers les différents domaines de recherche et applications.

La figure 2 est divisée en quatre quadrants, chacun représentant un domaine de recherche, en utilisant les mêmes couleurs distinctes que la Figure 1. À la différence de précédemment, je ne présente pas d'intersections entre les domaines mais une évolution dans le temps (cercles concentriques) du plus ancien (<2020) au plus récent (2024-2025). Les applications (en blanc) traversent à la fois cette chronologie et les domaines correspondants pour lesquels il y a eu une contribution significative. Chaque application est accompagnée d'une tâche spécifique (en vert) représentative des travaux que j'ai effectués. Enfin, chaque application qui a nécessité de considérer un ou plusieurs domaines de recherche, repose sur l'emploi de méthodes spécifiques à ces domaines (en noir).

Dans chaque domaine de recherche, les contributions se sont concentrées soit sur l'amélioration des performances de méthodes existantes, soit sur la création de nouvelles approches plus adaptables aux particularités des applications concernées, qui présentent leurs propres défis et contraintes liés aux données. Par exemple, dans les systèmes de recommandation, on doit gérer des contraintes comme la non-stationnarité (évolution de la probabilité de récompense au fil du temps) et la parcimonie vectorielle (concept-shift ou retours partiels des

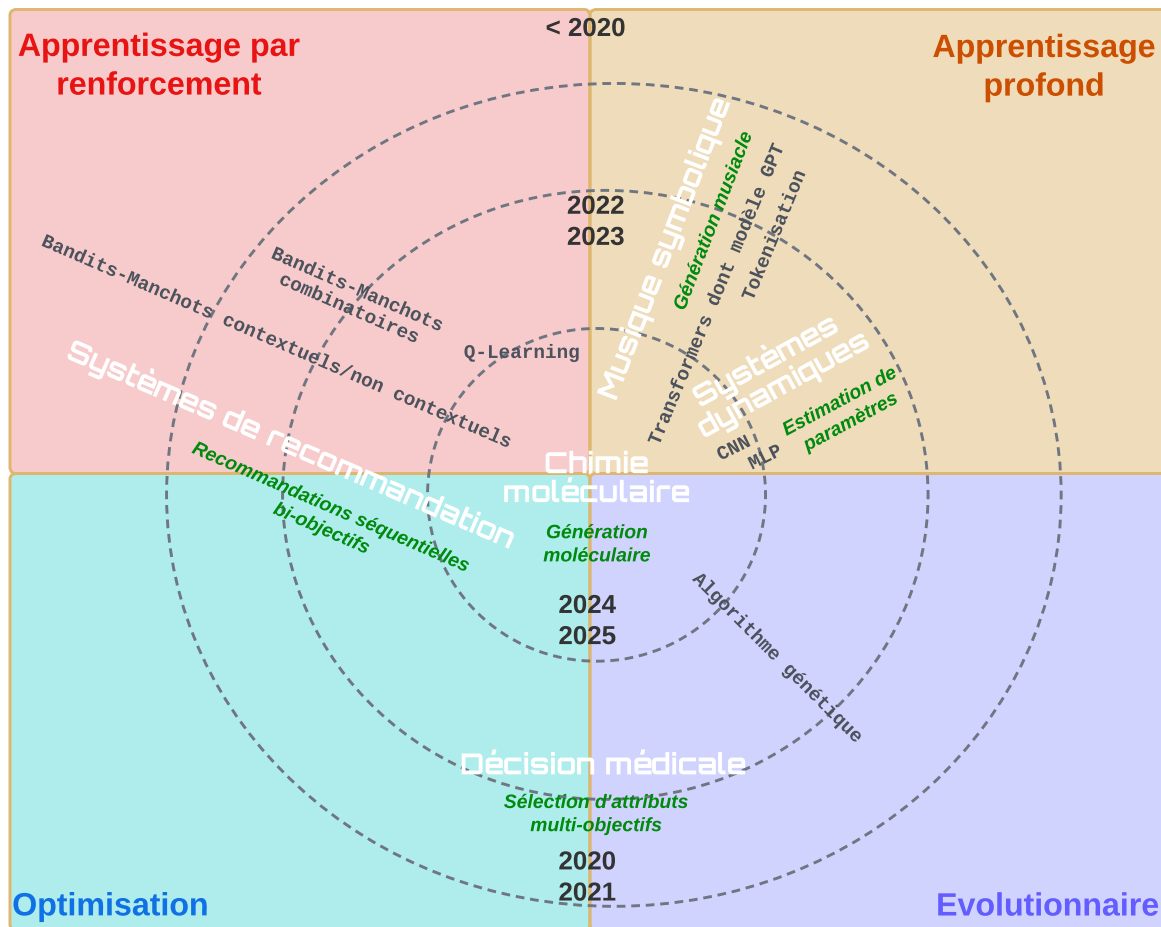


FIGURE 2 – Articulation chronologique des travaux de recherche à travers les différents domaines de recherche et applications

utilisateurs). En ce qui concerne les décisions médicales, il est crucial de sélectionner les attributs pertinents tout en maintenant une haute qualité de classification afin de réduire la dimension des variables décisionnelles. Pour l'estimation des paramètres dans les systèmes dynamiques, il est essentiel de fournir des approximations en temps réel tout en assurant la meilleure estimation possible par rapport aux méthodes de pointe actuelles, notamment grâce à des méthodes hybrides (basées sur des modèles et sur des données, c.-à-d., model-based et data-based) d'IA informée par les modèles (ou IA informée par la physique). Enfin, dans le domaine des méthodes de génération, il est question à la fois d'améliorer les techniques de tokenisation ou de résoudre le problème de la malédiction de la dimensionnalité dans certains cas d'étude pour améliorer la tâche d'apprentissage et de génération.

En reposant sur les Figures 1 et 2 présentées précédemment, je vais maintenant décrire plus précisément les différents travaux dans lesquels je me suis engagé pour chacune des

applications associées à leur chronologie respective et aux travers des quatre domaines de recherche clés.

Systemes de recommandation (2020-2024)

À la suite de ma thèse (soutenue le 04/11/2019), portant sur les systèmes de recommandation utilisant les algorithmes de bandits-manchots, j'ai poursuivi ma première année post-doctorale au sein de l'ESEO. Cette période m'a permis d'approfondir mes travaux et de les valoriser à travers une publication dans le journal *Information Sciences (Elsevier)*. Cette publication détaillait Gorthaur-EXP3 [Gut+21], une méthode de sélection dynamique d'algorithmes de recommandation basée sur les bandits-manchots, visant à optimiser les performances en temps-réel dans un contexte multicritère, incluant la précision et la diversité, tout en s'adaptant à des contraintes de non-stationnarité ou de parcimonie vectorielle.

En parallèle, j'ai co-encadré la thèse d'Alexandre Letard (2020-2023) [Let23], qui s'inscrivait dans la continuité de mes travaux. Son sujet portait sur un système d'apprentissage artificiel multicritère orienté par les besoins des utilisateurs. Plus précisément, cette thèse explorait l'utilisation de systèmes de recommandation pour identifier un sous-ensemble pertinent d'éléments à recommander parmi un vaste choix, en s'appuyant sur les algorithmes de bandits-manchots et l'optimisation multicritère. Nos contributions ont inclus l'analyse de l'impact des fonctions de récompense sur les performances de ces algorithmes, la formalisation d'un modèle générique pour ces fonctions, et la proposition de nouvelles méthodes pour la sélection des retours utilisateurs et le calcul des récompenses [Let+20b ; Let+22].

À la suite de cette thèse, nous avons continué nos travaux avec Alexandre Letard en adaptant l'algorithme Gorthaur-EXP3 [Gut+21] pour la recommandation de k éléments, en développant Multiple-Play Gorthaur et Budgeted-Gorthaur [Let+24]. Ces algorithmes ont pour objectif de satisfaire à la fois les critères de précision et de diversité des recommandations.

Ces travaux ont permis de significativement améliorer l'état de l'art dans le domaine des systèmes de recommandation multicritères, en offrant des solutions dynamiques et adaptatives aux défis posés par les données contextuelles complexes et la non-stationnarité.

Cet axe applicatif focalisé sur les systèmes de recommandation m'a permis d'aborder des problématiques en apprentissage par renforcement. Ces problématiques ont donné lieu à des contributions mêlant pratique et théorique, en particulier sur les problèmes de bandits-manchots, en ligne avec la thématique ARC du LERIA. La nécessité d'hybrider des méthodes m'a également conduit à explorer des problématiques liées à l'optimisation multicritère, en ligne avec la thématique MOC du LERIA. Ces expériences m'ont octroyé une vision transversale des enjeux posés par les systèmes de recommandation actuels, notamment en matière d'adaptation et d'optimisation dynamique.

Application aux systèmes de recommandation : principaux éléments et correspondance aux Figures 1 et 2

- Période : de 2020 à 2024
- Domaines de recherches abordés : Apprentissage par renforcement, Optimisation multicritère
- Méthodes employées : Algorithmes de bandits-manchots (combinatoires et à tirage simple)
- Tâches réalisées : Recommandations et top-k recommandations séquentielles bi-objectifs

Décision médicale (2021-2022) :

Mes travaux sur l'optimisation bi-objectif appliquée à l'apprentissage dans le contexte des systèmes de recommandation m'ont conduit à enrichir une méthode évolutionnaire développée pour la sélection d'attributs dans le cadre de l'aide à la décision médicale.

En 2021, j'ai poursuivi mes activités de recherche en collaboration avec le CHU d'Angers (Pr Pierre Abraham) et des chercheurs associés au LERIA (Dr Daniel Schang, ESEO-LERIA, et Dr Olivier Camp, ESEO-LERIA). Ce travail, dans la continuité de mes précédents travaux post-doctoraux (voir paragraphe précédent), portait principalement sur l'élaboration d'un « *Compass* » permettant le calcul de fitness multicritère et ce dans un cadre d'optimisation multicritère.

L'objectif de cette recherche était de développer une nouvelle méthode, GAWC [Gut+22] (Genetic Algorithm with multi-objective Compass), de sélection multicritère d'attributs pour les tâches de classification binaire. Celle-ci reposait sur :

1. Une approche ELM (Extreme Learning Machine), un type de réseau de neurones se distinguant par sa vitesse d'apprentissage extrêmement rapide et sa capacité à fournir une excellente performance de généralisation ;
2. Un algorithme génétique et un système de calcul de fitness innovant afin de sélectionner les variables de décisions médicales les plus pertinentes parmi un grand ensemble (par exemple, dans le cas du CHU, plus de 48 variables de décision) ;

La difficulté principale de cette sélection d'attributs résidait dans le maintien d'un bon niveau d'*AUC (Area Under the Curve)* et d'*exactitude (accuracy)* de classification, en évitant les erreurs de classification (faux positifs ou faux négatifs) tout en conservant un haut niveau de précision avec moins de variables de décision.

Notre nouvel algorithme s'est révélé supérieur aux méthodes concurrentes identifiées au niveau international sur plusieurs jeux de données réels (en plus de celui du CHU). Ce succès a souligné l'efficacité de notre approche combinant ELM et algorithme génétique pour l'optimisation multicritère dans le domaine médical.

Application à l'aide à la décision médicale : principaux éléments et correspondance aux Figures 1 et 2

- Période : de 2021-2022
- Domaines de recherches abordés : Optimisation multicritère, méthodes évolutionnaires
- Méthode employée : Algorithme génétique avec fonction de fitness améliorée
- Tâche réalisée : Sélection d'attributs multi-objectif (MOFS : Multi-Objective Features Selection)

Musique symbolique (2021-2024) :

En parallèle de la poursuite de mes travaux en apprentissage par renforcement et en optimisation multicritère, je me suis tourné vers l'étude des méthodes génératives reposant sur l'apprentissage profond, et plus particulièrement les *Transformers*. Cette nouvelle orientation, préparée en 2020 et lancée en 2021, a été motivée par trois facteurs principaux :

1. L'intuition depuis 2020 que les modèles génératifs basés sur les *Transformers* constituaient une innovation de rupture ;
2. Mon intérêt pour les méthodes génératives, formulé dans mon dossier de candidature en tant que projet d'intégration au LERIA en 2020 ;
3. L'expertise de longue date de mon laboratoire dans les techniques avancées de *TALN* (Traitement Automatique du Langage Naturel, c.-à-d., *Natural Language Processing : NLP*).

Concernant le premier facteur, la mise à disposition du grand public de ChatGPT fin novembre 2022 a largement dépassé mes attentes initiales, confirmant ainsi la révolution apportée par ces modèles.

En ce qui concerne le second point, après avoir encadré un stage de fin d'études (ESEO) en 2020, l'entreprise Aubay m'a manifesté son intérêt à financer une thèse CIFRE sur l'apprentissage profond appliqué à la génération de musique symbolique. Cette opportunité m'a permis d'associer mon intérêt pour les méthodes génératives et les besoins industriels, ouvrant ainsi la voie à une collaboration fructueuse et à des applications concrètes de mes recherches.

C'est donc dès 2021, au travers du spectre de la génération de musique symbolique, que j'ai découvert toutes les aptitudes et les promesses des modèles GPT en travaillant plus particulièrement sur les notions de tokenisation de la musique symbolique [Fra+21] et ses impacts dans la génération [Fra+23b]. Ces travaux ont été réalisés via le co-encadrement de la thèse de Nathan Fradet [Fra24], que j'ai co-montée avec Sorbonne Université (Dr Jean-Pierre Briot DR CNRS LIP 6, Pr Amal El-Fallah, LIP 6). Cette thèse a abouti à l'élaboration de techniques avancées de tokenisation, où nous avons notamment proposé l'utilisation du Byte Pair Encoding [Fra+23a] pour la tokenisation en musique symbolique.

Application à la musique symbolique : principaux éléments et correspondance aux Figure 1 et 2

- Période : de 2021-2024
- Domaine de recherche abordé : Apprentissage profond
- Méthodes employées : *Transformers*, *GPT-2*, *BERT*, Tokenisations
- Tâche réalisée : Génération

Systemes dynamiques (2022-2025) :

Mon intégration à l'IUT GEII en 2020 m'a permis de côtoyer des chercheurs de la section CNU 61 dont les domaines de recherche se rapprochaient de l'apprentissage automatique appliqué à la physique ou encore à la robotique. De même, les enseignements et projets étudiants en GEII se rapprochaient de la notion de capteur, de signal, de traitement du signal et de l'interprétation de la mesure physique.

C'est pourquoi, afin de me rapprocher des thématiques de ma composante, j'ai décidé à partir de 2021 de m'intéresser à celles-ci en parallèle de mes activités de recherche en apprentissage.

Pour ce faire, j'ai effectué une première tentative de réponse à un appel à projet (AAP) en 2021 pour une demande de financement de thèse AID (Agence Innovation Défense). Cette première tentative ayant échoué, je me suis appuyé sur les conseils du Pr Philippe Martinet (DR INRIA) pour revoir cette proposition en 2022. Après cette seconde tentative, le projet ayant été retenu, je suis devenu co-encadrant de la thèse à hauteur de 40% à partir d'octobre 2022. Cette thèse s'est déroulée sous la direction de Philippe Martinet (INRIA Sophia-Antipolis) en collaboration avec Sébastien Aubin de la Direction Générale de l'Armement (DGA). Le but principal de ce projet de recherche était de développer une méthodologie de conception d'un système d'Intelligence Artificielle hybride mêlant les approches basées sur des modèles (*model-based*) avec celles basées sur les données (*data-based*) pour l'évaluation du comportement dynamique (confort, sécurité, performances) de systèmes mobiles habités. L'application qui intéressait la DGA était l'analyse de la sécurité du comportement dynamique de véhicules militaires.

Les premiers travaux que nous avons menés avec notre doctorant Fabien Lioni se sont concentrés sur les méthodes bayésiennes pour l'estimation des paramètres de la dynamique latérale des véhicules terrestres [Lio+23 ; Lio+24]. Nous avons ensuite élargi ces travaux en développant une approche d'apprentissage par transfert, qui a permis de transférer les connaissances acquises sur un modèle spécifique, à d'autres configurations similaires ou connexes. Cette approche a montré des performances supérieures aux méthodes de référence dans la littérature.

Dans une seconde phase, nous avons conçu une méthode fondée sur l'estimation à horizon glissant (*MHE*), qui permet d'apprendre une relation directe entre les mesures du système et l'estimation d'états. Cette méthode repose sur des réseaux de neurones profonds (tels que *CNN*, *GRU* ou *CNN-T*), où la minimisation de la fonction de perte spécifique au MHE permet

de rendre les estimations plus rapides et précises, sans avoir besoin d'une optimisation en ligne.

Application aux systèmes dynamiques : principaux éléments et correspondance aux Figure 1 et 2

- Période : de 2022-2025
- Domaine de recherche abordés : Apprentissage automatique, apprentissage profond
- Méthodes employées : *CNN, GRU, CNN-T, MLP, ABC-SMC*
- Tâche réalisée : Estimation de paramètres et estimation d'état

Chimie moléculaire (2024-2025) :

Dans les paragraphes précédents, nous avons abordé différents projets et applications faisant appel à diverses compétences dans les quatre domaines de recherche que sont l'apprentissage par renforcement, l'optimisation, l'apprentissage automatique et profond, ainsi que les méthodes évolutionnaires. Ce paragraphe présente ma plus récente implication dans le domaine applicatif de la chimie moléculaire, en particulier dans la génération. Cette thématique représente non seulement la convergence de toutes les compétences que j'ai acquises précédemment en un seul projet, mais également les perspectives de mes recherches dans les méthodes génératives.

Lors de mon recrutement en 2020, mon projet d'intégration au LERIA visait à poursuivre mes travaux en apprentissage par renforcement, en les appliquant entre autres aux méthodes génératives. Mon objectif était de renforcer les compétences de l'équipe en y apportant mes connaissances en apprentissage par renforcement, tout en prenant en compte la notion de diversité dans ces méthodes.

Ainsi, au début de l'année 2024, j'ai obtenu des financements pour mon projet de recherche, intitulé ^G-*GENOCOD* (*Graph-Generation for Novel Compound Discovery*). Ce projet est soutenu par la région Pays de la Loire dans le cadre du programme « Étoiles Montantes », qui vise à identifier et accompagner de jeunes chercheurs prometteurs en leur permettant de développer leurs projets. Bien que ce programme récompense un seul lauréat, le projet s'appuie sur un consortium pluridisciplinaire de chercheurs. Il comprend le Dr. Benoit Da Mota et le Dr. Thomas Cauchy à l'origine des premiers travaux de ce champ de recherche (notamment *EvoMol* [Leg+20]), et le Pr. Sylvain Lamprier (expert en NLP). Les objectifs du projet ^G-*GENOCOD* comprennent le développement d'un module pour capturer les variables contextuelles moléculaires, l'implémentation de nouvelles méthodes d'apprentissage par renforcement couplées à *EvoMol*, ainsi que la création d'une application innovante pour la synthèse moléculaire basée sur des réactions connues. En parallèle, une thèse sur la même problématique, mais abordée sous l'angle de l'apprentissage profond et intégrant l'utilisation de *Transformers*, a débuté fin 2023. Cette thèse est dirigée par Sylvain Lamprier et co-encadrée par Benoit Da Mota et moi-même.

Application à la chimie moléculaire : principaux éléments et correspondance aux Figure 1 et 2

- Période : de 2024-2025
- Applications visées : Chimie moléculaire
- Domaine de recherche abordés : Apprentissage apprentissage profond, méthodes évolutionnaires, apprentissage par renforcement, optimisation
- Méthodes employées : *Q-Learning*, Bandits-manchots, algorithme évolutionnaire, *Transformers*, *GPT*
- Tâche réalisée : Génération

Rayonnement, animation et responsabilités scientifiques

Rayonnement et animation scientifique

Comité de Suivi Individuel de thèse

De 2021 à 2024, j'ai été membre du Comité de suivi individuel de la thèse de Nicolas Hubert dirigée par Davy Monticolo et Armelle Brun du laboratoire LORIA de l'université de Lorraine (Nancy). Son sujet portait sur la fouille de données, l'apprentissage statistique et les systèmes de recommandation pour l'éducation pour la conception d'un outil d'aide à la décision pour l'orientation scolaire.

Jury de thèse

Examinateur de la thèse de Zhao Zhang, soutenue le 5 Juillet 2022, et intitulée : « Recommandation de chemins d'apprentissage : un processus de décision séquentiel. »

- Thèse délivrée par l'Université de Lorraine
- Jury : Armelle Brun (Co-directrice de thèse ; LORIA - Université de Lorraine), Anne Boyer (co-directrice de thèse ; LORIA - Université de Lorraine), Marie-Hélène Abel (rapporteuse ; UMR CNRS 7253 Heudiasyc - UTC), Sylvie Calabretto (rapporteuse ; LIRIS - INSA Lyon), Davy Monticolo (examinateur - Président ; LORIA - ENSGSI Université de Lorraine), Nicolas Gutowski (examinateur ; LERIA - IUT Université d'Angers).

Expertises

- Je suis membre du Comité de programme de la conférence UAI depuis 2020 (5 articles par an à évaluer) et de la conférence AAAI depuis 2022 (6 articles à évaluer par an : 3 articles AAAI et 3 articles NIPS retransmis à AAAI). Ces deux conférences sont considérées comme les « top-ranked » des conférences en Intelligence Artificielle dans le monde. Depuis 2024 je suis relecteur pour ACL Rolling Review. ACL Rolling Review est

- une initiative visant à moderniser et améliorer le processus de soumission et de révision des articles de recherche en linguistique informatique, en offrant plus de flexibilité et de continuité dans la révision par les pairs. ACL Rolling Review est conçu pour des publications qui peuvent ensuite être présentées dans des conférences internationales de premier rang, souvent classées A* dans le domaine de la linguistique informatique et du traitement automatique des langues (c.-à-d., ACL, EMNLP, NAACL) ;
- De 2020 à 2024, j'ai été relecteur pour les revues internationales suivantes :
 - Journal of Information Sciences : 2 articles en 2022 et 1 article en 2024 (Journal Q1 en intelligence artificielle d'après Scimago Journal & Country Rank) ;
 - RAIRO-Operations Research journal : 3 articles depuis 2022 (Journal Q2 en Decision Sciences d'après Scimago Journal & Country Rank) ;
 - Journal of Expert Systems With Applications : 4 articles en 2023 et 2 articles en 2024 (Journal Q1 en intelligence artificielle d'après Scimago Journal & Country Rank) ;
 - J'ai été évaluateur pour une expertise de projet ANR (via la plate-forme IRIS) en 2022 ;
 - Je suis membre de l'Association française en Intelligence Artificielle (AfIA) depuis 2018 et de Madics (Masses de Données, Informations et Connaissances en Sciences) depuis 2016.

Organisation de conférences nationales et séminaires

- 2020 : Membre du comité d'organisation de la **Plate-Forme Intelligence Artificielle**, qui s'est tenue du 29/06/2020 au 03/07/2020. PFIA est portée par l'Association française pour l'Intelligence Artificielle (AfIA) et fût co-organisée cette année par le LERIA (Université d'Angers) et l'ESEO. Cette manifestation regroupe chaque année cinq conférences scientifiques majeures de l'intelligence artificielle en France (APIA, CNIA, RJCIA, JC, JIAF).
- 2022 : organisation d'un atelier TANGRAM à l'université d'Angers, en collaboration avec l'ingénieur filière Numérique & IoT Yves-Pol Cabon. Cet atelier à destination à la fois des académiques et des entreprises avait pour objectif de réaliser un retour d'expériences sur les différents montages et encadrements de thèse CIFRE que ce soit en termes administratif, organisationnel mais aussi scientifique.
- 2023 : organisation d'un séminaire à l'Université d'Angers, en collaboration avec Sorbonne Université et l'ESEO, portant sur l'IA et la Musique : https://leria.univ-angers.fr/?page_id=306. L'objectif de ce séminaire scientifique était de regrouper à la fois experts académiques et industriels du domaine de l'IA appliquée à la musique tout en en profitant pour présenter les travaux de la thèse que je co-encadrais en génération de musique. Ainsi, ce séminaire d'une matinée a pu faire intervenir le Pr Louis Bigo du LaBRI (Université de Bordeaux) et Groupement d'Intérêt Scientifique SCRIME ; le Dr Romain Hennequin, Responsable scientifique chez Deezer Research ; le Pr Jean-Pierre Briot Directeur de Recherche émérite CNRS LIP6 ; Nathan Fradet, Doctorant Sorbonne Université que je co-encadrais et aujourd'hui diplômé.

Invitation à des séminaires

Je suis intervenu dans le cadre du séminaire « Informatique et LLSHS : regards croisés » le 8 avril 2022. Durant cette journée organisée par la SFR Math-STIC et Confluences (Recherches pluridisciplinaires en lettres, langues, sciences humaines et sociales), j'ai abordé le sujet des systèmes de recommandation et plus particulièrement des biais. Cet événement local d'une cinquantaine de participants a fait intervenir sept enseignants-chercheurs de différentes disciplines autour de sujets divers au croisement de l'intelligence artificielle au sens large, de l'économie et des sciences humaines et sociales.

Cette présentation faisait suite à celle que j'ai effectué le 4 février 2022 dans le cadre d'un séminaire de la SFR Math-STIC où j'ai abordé la « confiance et les algorithmes de recommandation ».

J'ai également effectué deux présentations de vulgarisation en 2023 dans le cadre de séminaires organisés par Angers Technopole et DIVA (Digital Innovation Value Accelerator¹). DIVA est un hub financé par la Commission Européenne et la Région des Pays de la Loire et piloté par Atlanpole et le pôle de compétitivité EMC2. Ils ont fait appel à mes compétences via Angers Technopole pour ces séminaires avec pour objectif de démystifier l'IA pour le grand public et les décideurs d'entreprise. Ainsi durant celui du 13 avril 2023, j'ai expliqué les bases, évolutions et applications de l'IA et durant celui du 17 octobre 2023, j'ai présenté mes recherches appliquées, offrant une vue pratique de l'utilisation de l'IA dans divers secteurs.

Responsabilités scientifiques

- De 2020 à 2022 : En tant que Maître de Conférences depuis 2020, je me suis associé à des chercheurs expérimentés pour le montage, l'encadrement de projets de recherche et de thèses afin de bénéficier de leur expérience de direction doctorale et de réponse d'appel à projets. Ainsi, entre 2020 et 2022 j'ai porté en tant que co-responsable scientifique l'élaboration de quatre projets de thèse dont deux ont été retenus et ainsi financés : **une thèse CIFRE en 2021** pour l'entreprise Aubay avec le Pr Amal El-Fallah et le Pr Jean-Pierre Briot ; une thèse de l'Agence Innovation Défense (AID) du Ministère des Armées avec le Directeur de Recherche Philippe Martinet (INRIA Sophia-Antipolis et le Dr Sébastien Aubin.
- 2023 : J'ai été porteur d'un dossier de réponse à **AAP Étoiles Montantes de la région Pays de la Loire** qui a été accepté (démarrage le 1er janvier 2024). Ce projet : ^{G-}GENOCOD a été monté en collaboration étroite avec Benoit Da Mota (LERIA) positionné avec moi comme co-responsable scientifique, Sylvain Lamprier (LERIA) et Thomas Cauchy (MOLTECH-ANJOU). ^{G-}GENOCOD propose une approche novatrice en combinant l'apprentissage par renforcement (RL : Reinforcement Learning) avec un algorithme évolutionnaire pour générer des structures graphiques complexes, avec un accent particulier sur la génération moléculaire pour les matériaux.

1. <https://www.edih-diva.eu/>

- De 2024 à 2025 : Fort de deux thèses co-encadrées et soutenues en mai 2023 et en mars 2024, j'ai poursuivi mon encadrement de la thèse de Fabien Lioni qui a poursuivi sa seconde partie de thèse à Angers. J'ai également débuté ma mission de **coordinateur du projet** ^G**GENOCOD** comportant en plus des 4 enseignants-chercheurs titulaires à l'appui : un post-doctorant (18 mois) et un stagiaire M2 recherche (6 mois). ^G**GENOCOD** est fortement connexe à la thèse de Dérin Özer que je co-encadre depuis octobre 2023. Cette thèse porte sur la génération de structures graphiques appliquée à la chimie moléculaire.

Encadrements

Encadrements de thèse

Depuis ma soutenance de thèse le 04/11/2019, j'ai eu l'opportunité de co-encadrer **quatre thèses dont deux ont été soutenues en 2023 et en 2024**, et deux autres sont en cours.

De janvier 2020 à mai 2023 : Co-encadrement du doctorat d'Alexandre Letard à hauteur de 30% (ingénieur ESEO promotion 2018), avec Tassadit Amghar de l'université d'Angers et Olivier Camp de l'ESEO. Le sujet de la thèse était « Un système d'apprentissage artificiel multi-critères guidé par les besoins utilisateur ».

- Thèse soutenue en mai 2023 et délivrée par l'Université d'Angers.
- Rapporteurs : Pr. Davy Monticolo (Université de Nancy, LORIA), Pr. Anne Boyer (Université de Nancy, LORIA)
- Publications associées : [Let+21 ; Let+20a ; Let+20b ; Let+22]

D'avril 2021 à mars 2024 : Co-encadrement du doctorat de Nathan Fradet à hauteur de 30% (ingénieur ESEO promotion 2020), avec les Professeurs Jean-Pierre Briot et Amal EL-Fallah Seghrouchni de Sorbonne université (LIP 6) et Fabien Chhel de l'ESEO. Le sujet de la thèse était « Apprentissage automatique pour la génération conditionnée de musique symbolique multipiste. ».

- Thèse soutenue en mars 2024 et délivrée par Sorbonne Université.
- Rapporteurs : Pr. Louis Bigo (ENSEIRB-MATMECA, LABRI), Pr. Philippe Pasquier (Simon Fraser University)
- Publications associées : [Fra+21 ; Fra+23a ; Fra+23b]

Depuis septembre 2022 : Co-encadrement du doctorat de Fabien Lioni à hauteur de 40% (Master Data Science de l'Université Paris-Saclay et ingénieur ESME), avec le Professeur Philippe Martinet (DR) de l'Université de Côte d'Azur (INRIA Sophia-Antipolis. Le sujet de la thèse est « Évaluation de comportement dynamique par intelligence artificielle. Application à l'analyse de la sécurité du comportement dynamique d'un véhicule. ». Soutenance prévue pour

septembre 2025 (date ferme car fin de financement). Publications associées à date d'écriture de ce mémoire : [Lio+23 ; Lio+24] ainsi que deux articles acceptées en cours de publication² et une soumission d'article en conférence en cours de relecture.

Depuis octobre 2023 : Co-encadrement du doctorat de Dérin Özer (Master de Polytech Angers) à hauteur de 30%, avec le Professeur Sylvain Lamprier et le Dr Benoit Da Mota de l'Université d'Angers (LERIA). Il s'agit d'un financement académique de l'Université d'Angers. Le sujet de la thèse est « Génération de structures graphiques par apprentissage par renforcement profond : application à la chimie moléculaire ». À date d'écriture de ce mémoire, un article a été soumis en conférence et est en cours de relecture (**Prépublication** : <https://arxiv.org/html/2503.05810v1>).

Encadrement de post-doctorat

Depuis le 1^{er} septembre 2024 j'ai également l'opportunité d'encadrer une **post-doctorante**, Gaëlle Milon-Harnois, pour une durée de 18 mois. Cet encadrement fait suite à l'obtention du financement régional du dispositif Étoiles Montantes octroyé pour 2 ans (2024-2025) pour mon projet ^G GENOCOD.

Production scientifique

Mon apport à la recherche scientifique, depuis l'obtention de mon doctorat, est récapitulé dans le tableau 2 présenté ci-dessous. En se focalisant uniquement que sur les articles publiés dans des conférences de rang A ou supérieur, et dans des revues internationales classées dans le premier quartile (Q1), j'ai co-signé sept publications durant la période de mars 2020 à mars 2025, soit une moyenne de 1,4 articles par an. Parmi ces contributions, j'ai eu l'honneur d'être le premier auteur pour deux de ces articles (revues). Pour les autres, j'ai majoritairement endossé le rôle de second auteur en tant qu'encadrant de thèse, en me concentrant ainsi plus particulièrement sur la conception méthodologique, la validation des algorithmes et des modèles développés par mes doctorants, la participation à la rédaction des articles, leurs révisions successives si besoin et, pour l'un d'entre eux, la formalisation de la preuve (sketch).

Une vue d'ensemble de toutes mes publications est disponible sur mon *profil Google Scholar*, sur mon *profil ORCID*, ou encore sur mon site <https://ngutowski.fr>.

2. À la conférence IEEE International Conference on Robotics & Automation (ICRA) et dans le Journal of Intelligent Information Systems (Springer Nature).

Nature de la production	Nombre
Articles dans des revues internationales de premier quartile (Q1)	3
Articles dans des revues internationales de deuxième quartile (Q2)	1
Articles dans des conférences internationales de rang A*	2
Articles dans des conférences internationales de rang A	2
Articles dans des conférences internationales de rang B/C	4
Articles dans des conférences nationales	3
Logiciels	3

TABLE 2 – Synthèse de la production scientifique (période post-doctorale de 5 ans)

Journal à comité de lecture

1. F. Lioni, N. Gutowski, S. Aubin, P. Martinet. Physics-Guided Approach with Transfer Learning in Vehicle Lateral Dynamics³. *Journal of Intelligent Information Systems (Springer Nature)*, 2025.
Prépublication : <https://ngutowski.fr/article1.pdf>
2. [Let+24] A. Letard, N. Gutowski, O. Camp, T. Aghar. Bandit algorithms : A comprehensive review and their dynamic selection from a portfolio for multicriteria top-k recommendation. *Expert Systems with Applications (Elsevier)*, Volume 246, 15 July 2024
Lien web : <https://www.sciencedirect.com/science/article/abs/pii/S0957417424000162>
3. [Gut+22] N. Gutowski, D. Schang, O. Camp, P. Abraham. A novel multi-objective medical feature selection compass method for binary classification *Artificial Intelligence in Medicine (Elsevier)*, Volume 127, May 2022
Lien web : <https://www.sciencedirect.com/science/article/abs/pii/S0933365722000422>
4. [Gut+21] N. Gutowski, T. Amghar, O. Camp, F. Chhel. Gorthaur-EXP3 : Bandit-based selection from a portfolio of recommendation algorithms balancing the accuracy-diversity dilemma. *Information Sciences (Elsevier)*, Volume 546, 6 February 2021
Lien web : <https://www.sciencedirect.com/science/article/pii/S0020025520308744>.

Conférences internationales à comité de lecture

1. F. Lioni, N. Gutowski, S. Aubin, P. Martinet. Bias-Variance Analysis of Multi-Step Loss Functions for Dynamical System Identification⁴. *International Joint Conference on Neural Networks (IJCNN)*, Rome (Italie), 2025.
Prépublication : <https://ngutowski.fr/article3.pdf>
2. F. Lioni, N. Gutowski, S. Aubin, P. Martinet. Learning Direct Solution in Moving Horizon Estimation with Deep Learning Methods⁵. *IEEE International Conference on Robotics*

3. Cet article, accepté le 21/03/2025, sera publié courant 2025.

4. Cet article, accepté le 31/03/2025, sera disponible dans les actes de conférence d'IJCNN en juin 2025.

5. Cet article, accepté le 27/01/2025, sera disponible dans les les actes de conférence d'ICRA en mai 2025.

& Automation (ICRA), Atlanta (USA), 2025.

Prépublication : <https://ngutowski.fr/article2.pdf>

3. [Lio+24] F. Lioni, N. Gutowski, S. Aubin, P. Martinet. Bayesian Approach for Parameter Estimation in Vehicle Lateral Dynamics. *27th International Symposium on Methodologies for Intelligent Systems (ISMIS)*, Poitiers (FR), 2024.
Lien web : https://link.springer.com/chapter/10.1007/978-3-031-62700-2_22#citeas
4. [Fra+23a] N. Fradet, N. Gutowski, F. Chhel, JP. Briot. Byte Pair Encoding for Symbolic Music. *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Singapore, 2023.
Lien web : <https://aclanthology.org/2023.emnlp-main.123/>
5. [Fra+23b] N. Fradet, N. Gutowski, F. Chhel, JP. Briot. Impact of time and note duration tokenizations on deep learning symbolic music modeling. 2023. . *International Society for Music Information Retrieval Conference (ISMIR)*, Milan (IT), 2023.
Lien web : http://ismir2023program.ismir.net/poster_45.html
6. [Let+22] A. Letard, T. Amghar, O. Camp, N. Gutowski. COM-MABs : From Users' Feedback to Recommendation. *Florida Artificial Intelligence Research Society (FLAIRS)*, Hutchinson Island (USA), 2022.
Lien web : <https://journals.flvc.org/FLAIRS/article/view/130560>
7. [Fra+21] N. Fradet, JP. Briot, F. Chhel, A. El Fallah Seghrouchni, N. Gutowski. MidiTok : A Python package for MIDI file tokenization. 2021. *International Society for Music Information Retrieval Conference (ISMIR)*, Online, 2021.
Lien web : <https://archives.ismir.net/ismir2021/latebreaking/000005.pdf>
8. [Let+20b] A. Letard, T. Amghar, O. Camp, N. Gutowski. Partial Bandit and Semi-Bandit : Making the Most Out of Scarce Users' Feedback. *International Conference on Tools with Artificial Intelligence (ICTAI)*, Baltimore (USA), 2020.
Lien web : <https://ieeexplore.ieee.org/abstract/document/9288230>

Conférences nationales à comité de lecture

1. [Lio+23] F. Lioni, N. Gutowski, S. Aubin, P. Martinet. Approche Bayésienne pour l'Estimation des Paramètres de la Dynamique Latérale d'un Véhicule *Conference on Artificial Intelligence for Defense*, Rennes (FR), 2023.
Lien web : <https://caid-conference.eu/wp-content/uploads/2023/12/CAID2023-actes.pdf>
2. [Let+21] A. Letard, T. Amghar, O. Camp, N. Gutowski. Bandits-Manchots Combinatoires : du retour utilisateur à la recommandation. *Conférence Nationale en Intelligence Artificielle*, Bordeaux (FR), 2021.
Lien web : https://insei.hal.science/CNIA_2021/hal-03321206v1

3. [Let+20a] A. Letard, T. Amghar, O. Camp, N. Gutowski. Bandit et Semi-Bandit avec Retour Partiel : Une Stratégie d'Optimisation du Retour Utilisateur. *5ème conférence sur les Applications Pratiques de l'Intelligence Artificielle (APIA)*, Angers (FR), 2020.

Lien web : <https://hal.science/hal-02947326/>

Logiciels

J'ai au total contribué au développement de trois logiciels :

1. Entre 2020 et 2022, j'ai contribué au développement de *Algossim*, une bibliothèque intégrant les algorithmes de bandits manchots les plus populaires. Le projet est accessible sur <https://github.com/ngutowski/algossim> et est actuellement utilisé par un consortium international incluant des enseignants-chercheurs de l'ESEO, du LERIA et du Vietnam.
2. En 2022, en collaboration avec le CHU d'Angers, j'ai participé au développement de *GaWC* (Genetic Algorithm with multi-objective Compass), une nouvelle bibliothèque Python dédiée à la sélection multicritère de variables pour des décisions médicales. Le logiciel est partagé et accessible sur <https://github.com/ngutowski/gawc>.
3. Dans le cadre de la thèse CIFRE de Nathan Fradet sur la génération de musique symbolique avec l'intelligence artificielle, nous avons développé *MidiTok* en 2021 [Fra+21]. Cette bibliothèque⁶, qui réunit les sept principales méthodes de tokenisation mondiale dans le domaine, a reçu plus de 700 étoiles de la part des utilisateurs et est consultable à <https://github.com/Natooz/MidiTok>. *MidiTok* est plébiscitée par de nombreux chercheurs et enregistre environ 500 téléchargements hebdomadaires, avec plus de 133 000 téléchargements depuis sa création. L'article que nous avons publié à la conférence IS-MIR en novembre 2021 est déjà cité plus de 50 fois.

Ces trois bibliothèques sont distribuées sous la licence Open Source GPL-3.0.

Réflexion prospective

Dans les sections précédentes, j'ai présenté l'ensemble de mon implication scientifique décrite respectivement par mes activités de recherche, mes animations et responsabilités scientifiques, mes encadrements, et enfin mes productions scientifiques. Quatre domaines de recherche ont été mis en lumière : l'apprentissage par renforcement, l'apprentissage profond, l'optimisation et le domaine de l'évolutionnaire.

Pour mieux préciser mes projets de recherche à venir et ma perspective de carrière, voici la part que représentent chacun des domaines dans lesquels je travaille au regard de mon implication scientifique actuelle (Figure 3).

6. Toutes les informations chiffrées concernant les statistiques de *MidiTok* sont données à date d'écriture de ce mémoire.

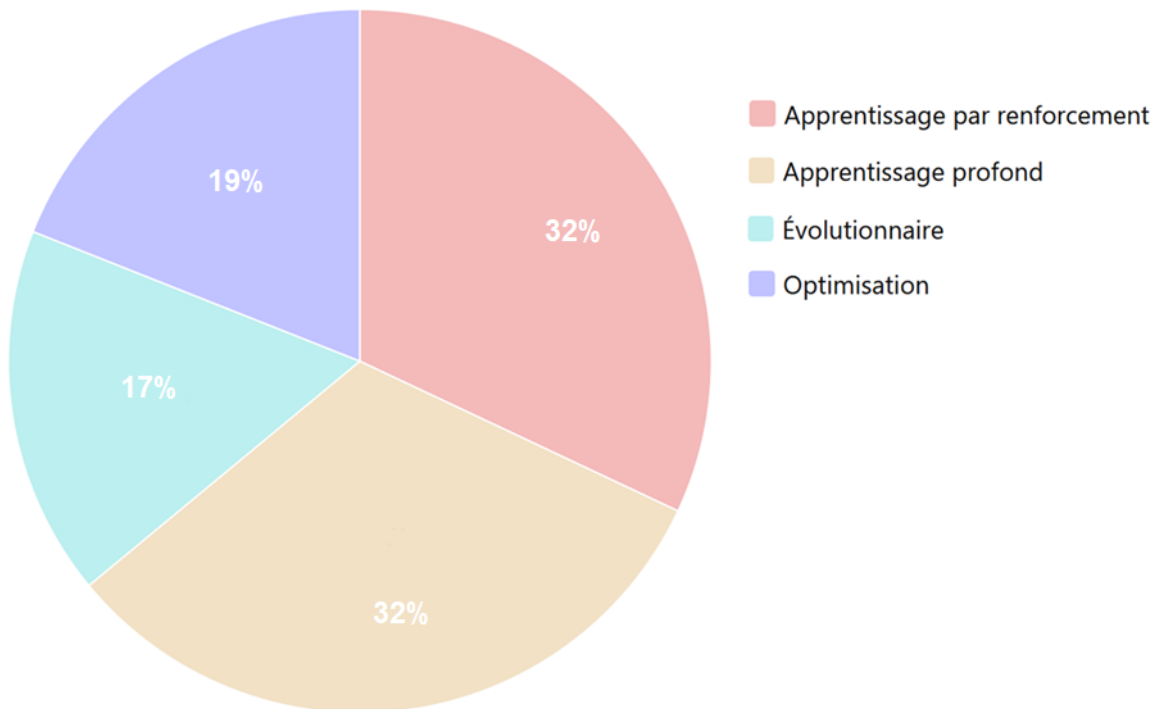


FIGURE 3 – Part de l'implication scientifique dans chaque domaine de recherche selon les publications et les encadrements

Pour ce faire, j'ai intégré les données de publication pondérées par leur niveau (journaux et quartile, conférences et rangs) et les données d'encadrements doctoraux et post-doctoraux pondérées par leur pourcentage.

On observe dans la Figure 3 que les deux domaines pour lesquels je me suis impliqué (apprentissage par renforcement, apprentissage profond) correspondant à la thématique ARC du LERIA représentent à eux seuls 64% de mon implication scientifique, tandis que l'optimisation et les méthodes évolutionnaires, relevant plutôt de la thématique MOC du LERIA, en représentent quant à eux 26%. Cette répartition souligne la logique de mon engagement scientifique, où l'apprentissage automatique d'ordre général constitue une part prépondérante (ou dominante) de mes contributions.

On note cependant que cette dominante se positionne très souvent à l'intersection des domaines de l'optimisation et des méthodes évolutionnaires. Cette hybridation a été nécessaire pour résoudre certains problèmes scientifiques spécifiques (par exemple, les systèmes dynamiques) ou certaines applications (par exemple, les systèmes de recommandation et la génération moléculaire).

En termes de projets de recherche à venir sur le long terme, j'envisage de poursuivre mes travaux sur les méthodes d'IA génératives appliquées (transdisciplinaire) et d'approfondir mes recherches dans le domaine de l'Intelligence Artificielle (IA) informée / guidée par la physique ou chimie. De ce fait, je commencerai dans un premier temps à déposer une réponse à appel à

projet ERC (octobre 2025) dans la continuité de mon projet ^{G-} *GENOCOD* du dispositif “Étoiles Montantes”.

De même, sur du moyen terme j’ai pour projet de co-monter une thèse portant sur l’optimisation et la modélisation de systèmes autosuffisants pour le traitement des eaux usées et la production d’énergie propre. Celle-ci portera donc sur la modélisation, la gestion de l’incertain et optimisation multicritère par intelligence artificielle en environnements non-linéaires à faible volume de données.

J’envisage également la poursuite des partenariats avec la DGA et l’INRIA sur les sujets portants sur l’IA informée par la physique.

L’ensemble de ces projets donne une vision allant jusqu’à 5 ans (voire plus en cas de succès d’ERC et de pérennisation des partenariats avec les institutions/entreprises actuelles) à compter de la publication de ce mémoire. Ceci me permet ainsi de continuer à poser le socle nécessaire à l’atteinte de mon objectif long terme, c.-à-d., approfondir mes recherches pour l’amélioration de la performance des méthodes d’IA générative appliquée et des méthodes d’IA informée / guidée par les modèles.

PREMIÈRE PARTIE

Optimisation multicritère appliquée à l'apprentissage

Préambule sur les termes multi-objectif et multicritère. Dans ce mémoire, les termes multi-objectifs et multicritères, bien que parfois proches, sont employés selon le type d'approche mentionnée. Nous partons du principe qu'une problématique multi-objectif vise à optimiser simultanément plusieurs objectifs distincts, ce qui conduit généralement à générer un ensemble de solutions représentant un compromis sur un front de Pareto. En revanche, nous considérons qu'une démarche multicritère se concentre sur l'évaluation ou l'agrégation de plusieurs critères, souvent combinés en une métrique unique pour identifier une solution optimale ou préférée. La méthode principale que nous avons employée dans nos propositions du Chapitre 1 et 2 – le *Compass* [MS08] – se situe à la frontière entre les deux principes.

Plus concrètement, que ce soit pour la suite de méthodes *Gorthaur* [Gut+21 ; Let+24] ou la méthode *GAwC* [Gut+22] que nous avons développées, l'objectif principal était de maximiser différents objectifs, ce qui correspond typiquement à une problématique multi-objectifs. Cependant, la particularité du cœur de la méthode utilisée (*Compass*) dans la suite de méthodes *Gorthaur* et dans la méthode *GAwC* réside dans son mécanisme d'agrégation. En effet, les différents objectifs (p.ex., diversité, nombre de caractéristiques, *accuracy*, *AUC*) sont combinés en une seule métrique de fitness via un vecteur de référence à plusieurs dimensions, reflétant un compromis défini par l'utilisateur. Ce mécanisme introduit une dimension multicritère, car il permet de pondérer et d'agréger les objectifs pour guider la recherche vers une solution optimale unique selon les préférences de l'utilisateur.

De ce fait, la distinction entre multi-objectif et multicritère peut s'estomper dans ce type d'applications. Par exemple en sélection d'attributs, minimiser le nombre d'attributs tout en maximisant l'*accuracy* serait classiquement une approche multi-objectifs, tandis que maximiser une fonction combinant ces deux critères s'apparenterait davantage à une démarche multicritères. Cependant, ces définitions ne sont pas toujours rigides, et certaines méthodes, notamment celles utilisant des mécanismes d'agrégation des objectifs comme dans le cas du *Compass*, peuvent s'inscrire à la croisée de ces deux paradigmes.

Ainsi, bien que la suite de méthodes *Gorthaur* et la méthode *GAwC* relèvent principalement d'une approche multi-objectifs (par son exploration du front de Pareto), leur utilisation d'un *Compass* introduit une composante multicritère pour faciliter le choix et la priorisation des solutions en fonction des besoins spécifiques.

Par conséquent, nous préférons nous référer par la suite au terme « multicritère » quand nous évoquons spécifiquement nos applications, le fonctionnement de nos méthodes ainsi que leur type d'optimisation, et au terme « multi-objectif » quand il s'agira de formaliser le problème ou présenter certaines méthodes compétitrices de la phase expérimentale.

Problèmes d'optimisation multi-objectif. Les problèmes d'optimisation multi-objectif sont réputés pour leur complexité, principalement en raison de l'aspect subjectif de ce qui constitue une solution optimale [Bou20 ; Coe07]. Quand deux solutions optimisent chacune un critère différent, il peut être difficile de déterminer laquelle est la meilleure, cette évaluation dépendant souvent du point de vue du décideur p.ex., expert métier, médecin, utilisateur d'un système de recommandation. Cette complexité s'accroît particulièrement dans le contexte des pro-

blèmes combinatoires, qui sont souvent NP-difficiles, nécessitant ainsi des approches sophistiquées pour identifier des solutions équilibrées.

Une méthode courante consiste à agréger des objectifs multiples en un objectif unique (méthode de scalarisation), ce qui permet de transformer le problème multi-objectifs en un problème plus simple à objectif unique [Gun18 ; MA10 ; PS13 ; ZHO17]. Une autre approche consiste à utiliser une somme pondérée des performances [MA10 ; WOJ17] ou la scalarisation de Chebyshev [Mav09 ; ZHO17]. D'autres stratégies font appel à des algorithmes évolutionnaires qui exploitent des règles de dominance, telles que les dominances de Pareto et de Lorenz, pour naviguer dans l'espace des solutions possibles [Bou20 ; Coe99 ; Deb11 ; Gun18 ; Mou88 ; Pru21]. Ces méthodes, souvent basées sur des populations, sont efficaces pour explorer simultanément plusieurs solutions de qualité.

On retrouve des problèmes d'optimisation multi-objectifs dans de nombreuses applications comme p.ex., les systèmes de recommandation, la planification, la robotique, la sélection d'attributs, l'optimisation d'hyperparamètres, la sélection d'opérateurs [Deb11 ; Gut+19b ; Kar+15 ; Li+20 ; MS08].

Nos travaux se sont plus particulièrement focalisés sur les systèmes de recommandation reposant sur l'apprentissage par renforcement (bandits-manchots) d'une part, et sur la sélection d'attributs médicaux dans le cadre de la classification binaire d'autres parts. De plus, nous nous sommes plus particulièrement concentré sur une stratégie exploitant les règles de dominances de Pareto.

Contributions de la Partie I. Le sujet de cette Habilitation à Diriger des Recherches s'est plus particulièrement focalisée sur l'apprentissage automatique pour la génération et l'exploration avec un focus particulier sur l'adaptation et la contextualisation pour données complexes. Cette partie met l'accent sur l'optimisation multicritère appliquée à l'apprentissage, en présentant des travaux autour de deux cadres applicatifs différents :

1. **Les systèmes de recommandation multicritères** avec recommandations simples ou multiples (voir Chapitre 1 de cette partie). Ce premier chapitre cible l'adaptation face à un environnement dynamique et des données pouvant être incertaines, dans le cadre de décisions de recommandation bi-objectifs.
2. **La sélection multicritère d'attributs médicaux** (voir Chapitre 2 de cette partie). Le second chapitre, quant à lui, explore les problématiques liées au fléau de la dimensionnalité et au raisonnement sur des données de décision complexes, où il est nécessaire de réduire le nombre de dimensions par la sélection d'attributs tout en préservant la capacité de raisonnement sur ces données.

Organisation de la Partie I. La Partie I est organisée comme suit :

1. Chapitre 1 - Apprentissage par renforcement et systèmes de recommandation multicritères
2. Chapitre 2 - Optimisation multicritère pour la sélection d'attributs médicaux

APPRENTISSAGE PAR RENFORCEMENT ET SYSTÈMES DE RECOMMANDATION MULTICRITÈRES

About Multi-Objective Recommender Systems...

« *Beyond recommendation accuracy, a wider perspective towards the recommendation utility may include other quality metrics, such as novelty and diversity, serendipity, popularity, and so forth. Maximizing one of these metrics may hurt others. Therefore, there is a demand in multi-objective optimization so that recommendations can be produced by balancing these metrics.* » [ZW22]

Yong Zheng & David Wang - Neurocomputing, 2022

Sommaire

1.1 Concepts généraux sur les systèmes de recommandation et les algorithmes de bandits-manchots	48
1.2 Positionnement des contributions	55
1.3 Méta-sélection d'algorithmes d'apprentissage	57
1.4 Méthodes <i>Gorthaur</i> : Méta-apprentissage pour la recommandation	59

Ce chapitre a pour objectif

- de rappeler les principes et les approches des systèmes de recommandation, en mettant progressivement l'accent sur leur utilisation basée sur les modèles, dans un contexte d'optimisation multicritère. Nous nous concentrerons ensuite sur les méthodes adaptées à la parcimonie dans les données et aux environnements non-stationnaires ;
- de réaliser la synthèse de nos contributions sur les systèmes de recommandation et problèmes multi-objectifs en environnements dynamiques et non stationnaires, avec données incertaines ;
- de mettre en œuvre une discussion sur nos propres contributions au champ de recherche visé dans ce chapitre.

Contexte des contributions de ce chapitre

Historique. Ce chapitre et les travaux qui en résultent sont liés à mes recherches post-doctorales à l'ESEO ainsi qu'à la thèse CIFRE d'Alexandre Letard chez Kara Technology (2020-2023) [Let23]. Le sujet de cette thèse a émergé à la suite de mon encadrement de son projet de fin d'études chez SPIE, qui portait notamment sur l'apprentissage profond. Pour répondre à son souhait de réaliser une thèse industrielle, j'ai co-construit le sujet en collaboration avec le futur doctorant, l'entreprise, l'ESEO et le LERIA. Intitulée « *Un système d'apprentissage artificiel multicritère guidé par les besoins utilisateur* », cette thèse a été dirigée par Tassadit Amghar, et j'ai eu le plaisir de la co-encadrer de 2020 à 2023 avec Olivier Camp (ESEO).

Équipe. Les institutions impliquées dans les recherches présentées dans ce chapitre étaient l'ESEO, le LERIA et l'entreprise Kara Technology plus particulièrement dans le cadre de la CIFRE.

Environnement et organisation. Les travaux se sont en partie déroulés durant la pandémie de *COVID-19* impliquant un travail de recherche à la fois en distanciel et en présentiel. Concernant l'encadrement de la thèse, les réunions de suivi se déroulaient au fil de l'eau mais nous avons un point récurrent systématique au moins une fois par mois.

Mon implication et mes contributions. Pour la partie concernant mes travaux post-doctoraux, j'ai assuré les missions de formalisation, conception, implémentation, expérimentation et écriture/revue d'articles. Pour la partie concernant la thèse, j'ai été impliqué dans le co-encadrement à hauteur de 30%. Cette mission englobait :

- le suivi régulier du doctorant ;
- le conseil scientifique et stratégique ;
- la participation à l'écriture et la revue d'articles ;
- la participation aux conférences dans lesquelles nous avons publié ;
- l'aide à la constitution du jury de thèse.

Les travaux de cette thèse ont été ensuite poursuivis et étendus dans un article de revue pour lequel j'ai été impliqué sur : la conception de la méthode, la validation, l'écriture et la revue de l'article (incluant la preuve théorique – sketch – dans son intégralité).

Note importante. Les sections de ce chapitre s'appuient sur les travaux suivants : [Gut+19b ; Gut+21 ; Let23 ; Let+20b ; Let+22 ; Let+24].

Ces travaux ont été réalisés en collaboration avec les co-auteurs et les contributions intellectuelles sont ainsi partagées entre eux. De même, le mémoire de thèse d'Alexandre Letard résulte d'un travail conjoint entre le doctorant (producteur du mémoire) et ses co-encadrants (relecteurs et conseillers). Ainsi, certaines sections de ce chapitre peuvent provenir de traductions, de reformulations ou encore de formulations textuelles de ces travaux originaux.

1.1 Concepts généraux sur les systèmes de recommandation et les algorithmes de bandits-manchots

Cette section fait un rapide état de l’art sur les systèmes de recommandation et présente les définitions clés sur les bandits-manchots pour la recommandation. Pour ce faire, elle repose sur des éléments présentés dans [Gut19 ; Gut+21 ; Let23 ; Let+24].

1.1.1 Rappel sur les systèmes de recommandation

Les systèmes de recommandation assistent activement les utilisateurs dans leurs choix parmi diverses options [Bur02 ; RRS21]. Ces systèmes ont été largement étudiés [Ben17 ; Gut19 ; Gut+19c ; Lou16 ; Pic11] et ont évolué au fil du temps pour couvrir de nombreuses applications.

Pour répondre aux divers enjeux des systèmes de recommandation, plusieurs approches spécifiques ont été développées, chacune avec ses propres limitations et axes d’amélioration. N’existait pas encore de représentation unifiée qui les regroupe toutes de manière consensuelle, nous nous reposons sur la Figure 1.1 proposée dans [Gut19] et qui présente les principales approches identifiées à partir des travaux de [JF13 ; Lu+15 ; Neg15 ; RRS15 ; RRS21].

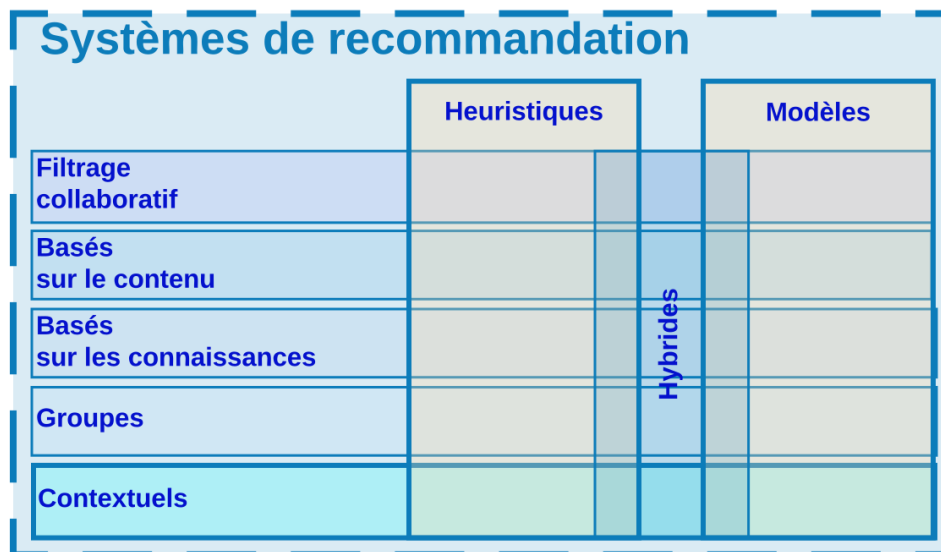


FIGURE 1.1 – Différentes approches utilisées par les systèmes de recommandation selon [Gut19]

À partir de ces études, cinq approches principales pour les systèmes de recommandation peuvent être dégagées [Gut19] :

1. Les systèmes de recommandation utilisant le filtrage collaboratif [RRS15 ; RRS21]. Le filtrage collaboratif (*Collaborative Filtering - CF*) permet de recommander un élément

- à une personne en se basant sur les opinions d'autres personnes ayant des centres d'intérêts similaires [DK04].
2. Les systèmes de recommandation basés sur le contenu (*Content-Based*) [RRS15 ; RRS21]. Les systèmes de recommandation basés sur le contenu (*Content-Based - CB*) recommandent des articles ou produits similaires aux éléments précédemment appréciés par un utilisateur [PB07].
 3. Les systèmes de recommandation fondés sur les connaissances (*Knowledge-Based*) [RRS15 ; RRS21]. Les systèmes de recommandation basés sur les connaissances (*Knowledge-Based - KB*) permettent la recommandation d'éléments à des utilisateurs à partir de la connaissance que le système possède sur les utilisateurs, les éléments à recommander, et la relation existante entre utilisateurs et éléments.
 4. Les systèmes de recommandation destinés à des groupes d'utilisateurs. Ce type de systèmes de recommandation est généralement utilisé pour produire des suggestions à des groupes d'utilisateurs lorsque les membres de ce groupe doivent trouver un consensus [McC+06 ; Oco+01 ; Pop13].
 5. Les systèmes de recommandation contextuels [Gut+17]. Ce type de système propose des suggestions personnalisées en tenant compte du contexte spécifique de l'utilisateur [Gut+19c], comme l'heure, la localisation, ou l'appareil utilisé. Il utilise des données contextuelles [AT11] en complément des préférences et comportements passés pour affiner la pertinence des recommandations.

Chaque approche peut être mise en œuvre de différentes manières, soit par des heuristiques, soit par des modèles utilisant des techniques d'Intelligence Computationnelle (*Computational Intelligence*). Par exemple, [RRS15 ; RRS21] distingue les approches basées sur des modèles de celles fondées sur des heuristiques.

Dans le cadre des approches basées sur les modèles, un domaine d'intérêt est celui des bandits-manchots, approches qui permettent un apprentissage par renforcement en équilibrant exploration et exploitation [LS20 ; Rob52 ; Tho33]. Ces méthodes ont été adaptées pour traiter des scénarios multi-objectifs [DN13 ; Gut19 ; Gut+19b ; Öne+18 ; RZN17] et pour aborder des problèmes tels que la sélection d'algorithmes [GS11 ; Gut+19b ; LF18]. Plusieurs travaux ont également exploré leur application à la recommandation d'éléments multiples, qu'il s'agisse de méthodes basées sur des modèles stochastiques [Aue02 ; SB98], bayésiens [AG12], ou d'algorithmes avec adversaires [Aue+00 ; Aue+02].

1.1.2 Rappel sur les bandits-manchots pour la recommandation

1.1.2.1 Définitions

Le problème du bandit-manchot (*Multi-Armed Bandit - MAB*) a été largement étudié depuis sa première formulation par [Rob52] en 1952.

Le problème peut être illustré simplement en considérant un joueur face à des machines à sous (bandits-manchots) dans un casino. Le but du joueur est de tirer à chaque fois sur le levier

(bras) le plus rémunérateur lorsqu'une pièce est insérée dans la machine. Le problème *MAB* fait référence à l'enjeu de développer une stratégie visant à déterminer, à chaque tour, un bras à tirer (sans aucune connaissance initiale du taux de gain de chaque bras) afin de maximiser le gain total. Il s'agit donc de trouver un compromis entre l'exploration nécessaire pour estimer la valeur de chaque bras, et l'exploitation qui consiste à se fier aux connaissances acquises à partir des expériences passées pour sélectionner les bras les plus rémunérateurs.

Une version étendue, connue sous le nom de problème de bandit-manchot contextuel (*Contextual Multi-Armed Bandit - CMAB*) [AG13b; LZ08], prend en compte des informations contextuelles. Plus précisément, dans les problèmes de *CMAB*, on considère que la récompense d'un bras dépend du contexte. L'enjeu devient donc, à chaque tour, de choisir le bras le plus rémunérateur en fonction du contexte.

Plusieurs solutions aux problèmes de *MAB* et de *CMAB* ont été développées : certaines utilisent des formulations stochastiques [Aue02; BF16], d'autres des approches bayésiennes [AG12]. Les problèmes de *MAB* et *CMAB* ont été adaptés pour répondre aux besoins des systèmes de recommandation. Ainsi, plusieurs travaux ont utilisé des *MABs* pour modéliser des problèmes de recommandation et expérimenté l'utilisation d'algorithmes tels que *UCB* [Aue02], *Thompson Sampling (TS)* [AG12] et *EXP3* [Aue+02] pour les résoudre. De même, les problèmes de recommandation contextuels, où la pertinence d'une recommandation dépend du contexte, ont été modélisés comme des *CMABs*, et des algorithmes tels que *LinUCB* [Li+10], *Linear Thompson Sampling (LinTS)* [AG13b] ont été largement expérimentés pour les résoudre.

Ci-dessous nous définissons les problèmes de *MAB* et de *CMAB* pour la recommandation ainsi que la notion de regret dans un tel cadre.

1.1.2.2 Bandits-manchots pour la recommandation : Définitions

En nous basant sur la définition originale du problème du bandit-manchot (*Multi-Armed Bandit - MAB*), nous définissons formellement [Gut+21] le problème de *MAB* pour la recommandation (Définition 1), les algorithmes qui résolvent ce problème (Définition 2) et la notion de regret (Définition 3).

Définition 1. Le problème du bandit-manchot (*Multi-Armed Bandit - MAB*) pour la recommandation. Soit $\mathcal{A} = \{a_1, \dots, a_m\}$ un ensemble de m bras indépendants à tirer. Dans le cas spécifique de la recommandation, il existe un élément à recommander correspondant à chaque bras. Soit D_r la distribution de l'espérance de récompense des éléments à recommander. Par conséquent, $D_r = (\mu_{a_1}, \dots, \mu_{a_m}) \in [0, 1]^m$, où $\mu_{a_i} \in [0, 1]$ est la probabilité de récompense lors de la recommandation de l'élément a_i , $i \in [1, m] \cap \mathbb{N}$. Le problème est séquentiel : à chaque itération $t \in [1, T]$, un utilisateur $u \in \mathcal{U}$ arrive et est pris en compte. Tout d'abord, un échantillon $(R_{a_1}, \dots, R_{a_m})$, $R_{a_i} \in \{0, 1\}$ est tiré de D_r . Ensuite, un élément $a_i \in \mathcal{A}$ est choisi par l'agent (le système de recommandation) et recommandé à u . Enfin, la récompense R_{a_i} de l'utilisateur u est révélée : 1 si l'utilisateur a apprécié l'élément recommandé et 0 sinon.

Définition 2. Algorithmes de MAB pour la recommandation. À chaque itération t , un algorithme de MAB b pour la recommandation détermine un élément $a_i \in \mathcal{A}$, $\mathcal{A} = \{a_1, \dots, a_m\}$ à recommander, en se basant sur la séquence des $t - 1$ observations précédentes $(a_{i,1}, R_{a_i,1}), \dots, (a_{i,t-1}, R_{a_i,t-1})$. Après avoir reçu la recommandation, l'utilisateur l'évalue et effectue un retour à l'agent (le système de recommandation) sous forme de succès (1) ou d'échec (0). À la réception du retour de l'utilisateur, l'agent met à jour le vecteur des récompenses \vec{R}_t .

L'objectif de l'agent est de maximiser la récompense totale attendue $\sum_{t=1}^T \mathbb{E}_{\vec{R}_t \sim D[R_{a,t}]}$. Une politique optimale connaît la récompense moyenne de chaque élément et recommande l'élément a^* ayant la récompense moyenne la plus élevée, c'est-à-dire $a^* = \arg \max_{a \in \mathcal{A}} (\mu_a)$.

Ainsi, afin de déterminer l'efficacité d'un algorithme de MAB b pour la recommandation, nous devons mesurer le regret cumulatif qu'il obtient $\rho_T(b)$ (où T est l'horizon) et le comparer à celui obtenu par la politique optimale. Nous pouvons donc définir le regret comme suit.

Définition 3. Regret dans les problèmes de MAB pour la recommandation. Dans le cas des systèmes de recommandation basés sur les bandits, un gain peut être considéré comme une recommandation réussie à un utilisateur u , et un regret comme un échec. Soit $g_T^* = \sum_{t=1}^T R_{a^*,t}$ le gain obtenu par une politique optimale à l'horizon T . Alors, l'espérance du gain de la politique optimale est $\mathbb{E}[g_T^*] = T \mu^*$. Soit b un algorithme de MAB pour la recommandation. Le regret cumulatif d'un algorithme ayant recommandé la séquence suivante d'éléments $a_{i,1}, \dots, a_{i,T}$ est donc $\rho_T = g_T^* - \sum_{t=1}^T R_{a_i,t}$.

1.1.2.3 Bandits-manchots contextuels pour la recommandation : Définitions

Le problème des bandits-manchots contextuels [Gut19 ; Li+10 ; NET20] (*Contextual Multi-Armed Bandit - CMAB*) est un cas particulier où l'on suppose une dépendance linéaire entre l'espérance de la récompense $\mathbb{E}[R_{t,a}]$ et le vecteur de contexte observé $x \subseteq \mathbb{R}^d$, de dimension d encodant les caractéristiques d'un utilisateur. Ainsi, l'espérance de la récompense est exprimée comme suit : $\mathbb{E}[R_{t,a}|x_t] = \hat{\theta}_{t,a}^\top x_t$, où $\hat{\theta}_{t,a}$ est un vecteur de coefficients associé au bras a , initialement nul et estimé à chaque itération t . Par conséquent, les définitions précédentes pour les problèmes de bandits-manchots sont modifiées dans le cas contextuel [Gut+21 ; LZ08], comme le montrent les définitions 4, 5 et 6.

Définition 4. Problèmes de bandit-manchot contextuel (Contextual Multi-Armed Bandit - CMAB) pour la recommandation. Soit $\mathcal{A} = \{a_1, \dots, a_m\}$ un ensemble de m bras indépendants à tirer, pour chacun desquels, dans le cas de la recommandation, il existe un élément à recommander. Soit la distribution $D_{x,R}$ sur $(x, R_{a_1}, \dots, R_{a_m})$. $D_{x,R}$ représente la distribution conjointe entre les contextes x et les récompenses R où $x \in X \cap \mathbb{R}^d$ est un contexte, et $R_{a_i} \in \{0, 1\}$ est la récompense associée à l'élément a_i , $i \in [1, m] \cap \mathbb{N}$. Le problème est séquentiel : à chaque itération t , un échantillon $(x, R_{a_1}, \dots, R_{a_m})$ est tiré de $D_{x,R}$, un utilisateur u avec son contexte x arrive et est observé, puis un élément a est sélectionné par l'agent, recommandé à u , et sa récompense R_a est révélée (1 si l'utilisateur a apprécié l'élément recommandé, 0 sinon).

Définition 5. Algorithmes de CMAB pour la recommandation. À chaque itération t , un algorithme de CMAB b détermine un élément à recommander $a_i \in \mathcal{A}$, $\mathcal{A} = \{a_1, \dots, a_m\}$. Ce choix est basé sur la séquence précédente d'observations $(x_1, a_{i,1}, R_{a_i,1}), \dots, (x_{t-1}, a_{i,t-1}, R_{a_i,t-1})$ et sur le contexte actuel observé x_t . L'algorithme met ensuite à jour le vecteur des récompenses \vec{R}_t selon le retour de l'utilisateur.

L'objectif est de maximiser la récompense totale espérée $\sum_{t=1}^T \mathbb{E}_{x, \vec{R}_t \sim D[R_{a_i,t}]}$. Soit $\Pi : X \rightarrow \mathcal{A}$ l'ensemble des politiques de recommandation possibles, où la politique optimale à déterminer est $\pi^* = \arg \max_{\pi \in \Pi} \mathbb{E}_{R,x}[R_{t,\pi(x)}]$. Ainsi, afin de déterminer l'efficacité d'un algorithme de CMAB b pour la recommandation, nous pouvons mesurer le regret cumulatif qu'il obtient $\rho_T(b)$ et le comparer à celui obtenu par la politique de recommandation optimale. Nous pouvons donc définir le regret comme suit.

Définition 6. Regret dans les problèmes de CMAB pour la recommandation. Dans le cas des systèmes de recommandation basés sur des bandits-manchots contextuels, une recommandation réussie donnée à un utilisateur u dans le contexte x est considérée comme un gain, tandis qu'une recommandation échouée est un regret. La récompense espérée pour une politique de recommandation $\pi \in \Pi$ est :

$$\mathcal{R}(\pi) = \mathbb{E}_{(x, \vec{R}) \sim D}[R_{\pi(x)}]$$

Considérons un algorithme de CMAB b . Soit $Z^T = \{(x_1, \vec{R}_1), \dots, (x_T, \vec{R}_T)\}$, et le regret espéré de b par rapport à la politique π est :

$$\Delta\rho(b, \pi, T) = T \mathcal{R}(\pi) - \mathbb{E}_{Z^T \sim D^T} \sum_{t=1}^T R_{b(x),t}$$

Le regret espéré de b jusqu'à l'horizon T par rapport à l'espace des politiques de recommandation Π est alors défini comme :

$$\Delta\rho(b, \Pi, T) = \sup_{\pi \in \Pi} \rho(b, \pi, T)$$

1.1.2.4 Limites des MAB/CMAB

Malgré les garanties théoriques des algorithmes de MAB et de CMAB, nous observons des résultats différents selon la nature des applications du monde réel ou des ensembles de données hors ligne [Gut+19b]. Ainsi, lorsqu'il s'agit de déployer un système de recommandation basé sur un modèle d'apprentissage automatique dans le monde réel, il est nécessaire de considérer au préalable quel algorithme pourrait le mieux répondre aux besoins de l'application. De plus, nous devons nous assurer que le système fournisse des recommandations à la fois précises et diversifiées.

Nos travaux se sont concentrés sur les algorithmes de bandits-manchots (non contextuel - MAB : *Multi-Armed Bandit*, et contextuels CMAB : *Contextual Multi-Armed Bandit*) utilisés

comme systèmes de recommandation bi-objectifs, que ce soit pour des tirages simples ou multiples [Gut+19b ; Gut+21 ; Let+24]. Nous avons abordé l'adaptation de ces algorithmes dans des environnements dynamiques et non stationnaires, ainsi que leur robustesse face à des données incertaines.

1.1.3 Rappel sur bandits-manchots pour la recommandation de plusieurs éléments

1.1.3.1 Bandits-manchots pour la recommandation de plusieurs éléments : Définitions

Le cadre des bandits-manchots pour la recommandation de plusieurs éléments inclut le problème bien connu des bandits-manchots Combinatoires (COM-MAB) [AVW87 ; Com+15], qui est une généralisation du problème de MAB où des ensembles de k bras, $S_t = \{a_1, \dots, a_k\}$, $S_t \subseteq \mathcal{A}$, avec $1 \leq k \leq m$, sont recommandés à chaque itération $t \in [1, T]$. Ces ensembles de bras sont généralement appelés "super-bras".

Les approches populaires pour résoudre les recommandations de plusieurs éléments impliquent soit l'utilisation de k instances d'un algorithme de *MAB/CMAB* [KSS ; RKJ08], l'exploitation d'un oracle d'approximation α [CWY13 ; NET20] ou l'utilisation directe de l'algorithme Multiple-Play [AVW87 ; Com+15 ; Ito+19] pour construire le super-bras. Dans le cadre de nos contributions, la méthode multiple-play sera utilisée pour permettre l'utilisation séquentielle des algorithmes de *MAB/CMAB* précédemment exposés pour construire itérativement S_t , comme suit : tant que $|S_t| < k$, $S_t = \cup_{i=1}^k a_i$ où $a_i = \operatorname{argmax}_{a \in \mathcal{A} \setminus S_t} \mathbb{E}[R_{t,a}]$. Le super-bras S_t est ainsi le sous-ensemble des k bras ayant les plus grandes attentes de récompense selon la politique de l'algorithme de MAB π . Cette généralisation présente plus de problématiques, tels que l'ordre des éléments ou la diversité individuelle. Comme déjà montré dans la littérature, identifier une combinaison optimale S^* est un problème NP-difficile [RKJ08].

Ainsi, les définitions pour les problèmes de *MAB* et *CMAB* données en Section 1.1.2.2 peuvent être ajustées pour la recommandation de plusieurs éléments comme suit [Let+24] :

Définition 7. Problème de bandit-manchot (Multi-Armed Bandit - MAB) avec plusieurs éléments. Dans un problème de MAB avec k éléments, il existe une distribution de récompense D_r avec une espérance $\mu_a \in [0, 1]$ et un support $R_{a_i} \in [0, 1]$ tel que $D_r = (\mu_{a_1}, \dots, \mu_{a_m}) \in [0, 1]^m$. Ici, $\mathcal{A} = \{a_1, \dots, a_m\}$ est un ensemble donné de m bras à tirer, et μ_{a_i} est la probabilité de recevoir une récompense pour le bras a_i ($i \in [1, m] \subseteq \mathbb{N}$). Ainsi, pour $k \leq m$, il existe C_m^k combinaisons possibles S de k bras. La valeur maximale de C_m^k est atteinte lorsque les m bras de \mathcal{A} sont indépendants, et $C_m^k = \frac{m!}{k!(m-k)!}$. De plus, il existe un ensemble de séquences de récompenses avec un support $R \in [0, 1]^{C_m^k}$, où R correspond au vecteur composé des récompenses retournées R_{a_i} pour chacun des bras sélectionnés, formant l'ensemble d'éléments S_t . Le problème est séquentiel : à chaque itération, un échantillon $(R_{a_1}, \dots, R_{a_m}), R_{a_i} \in \{0, 1\}$ est tiré de D_r . Ensuite, un agent choisit un ensemble S de bras $a_i \in \mathcal{A}$, et le vecteur de récompenses associé $R_S = (R_{a_1}, \dots, R_{a_i}, \dots, R_{a_k})$ est révélé. Étant donné que la distribution de récompenses D_r est

initialement inconnue, il est nécessaire de trouver un équilibre entre exploitation et exploration pour l'estimer correctement et maximiser les récompenses à long terme.

Définition 8. Algorithme de bandit-manchot (Multi-Armed Bandit - MAB) pour la recommandation à multiples éléments. Un algorithme de MAB b sélectionne un ensemble S de k bras $a_i \in \mathcal{A}$ ($a_i \in \mathcal{A} \setminus S$, dans le cas d'une sélection incrémentale), à chaque itération t , basé sur la séquence des observations précédentes $((S_1, R_{S,1}), \dots, (S_{t-1}, R_{S,t-1}))$. L'algorithme observe alors le vecteur de récompenses \vec{R}_t .

Définition 9. Regret. Le gain d'une politique optimale à l'horizon T est défini comme $g^T = \sum_{t=1}^T \sum_{j=1}^k R_{S_j,t}$. Soit S^* l'ensemble optimal de bras maximisant la récompense espérée. Ainsi, le gain espéré de la politique optimale est $\mathbb{E}[g^*(T)] = T \sum_{i=1}^k \mu_i^*$. Soit b un algorithme quelconque de bandit-manchot. Le regret cumulé de l'algorithme ayant joué la séquence de super-bras S_1, \dots, S_T est défini comme $\rho_T = \mathbb{E}[g^*(T)] - \sum_{t=1}^T \sum_{j=1}^k R_{S_j,t}$.

1.1.3.2 Fonctions de récompenses pour bandits-manchots à tirage multiple

La fonction de récompense pour bandits-manchots à tirage multiple diffère de la définition établie pour les bandits-manchots à tirage simple. Ainsi, nous rappelons ci-dessous son modèle en 3 étapes.

Modèle générique en trois étapes. Soit $\mathcal{U} = \{u_1, \dots, u_n\}$ un ensemble de n utilisateurs. À chaque itération $t \in [1, T]$, un utilisateur u_t , avec contexte x_t , demande une recommandation S_t de $k < |\mathcal{A}|$ éléments de $\mathcal{A} = \{a_1, \dots, a_m\}$. Un retour utilisateur $F_t \subseteq Y_{u,t}$ est alors émis, $Y_{u,t}$ modélisant l'opinion de u_t sur chaque $a \in \mathcal{A}$. La fonction de récompense $\phi(S_t, Y_t)$ permet de déterminer F_t et de calculer un vecteur de récompenses R_t pour mettre à jour la politique π de l'algorithme MAB.

Toutes les fonctions de récompense se basent sur trois processus successifs :

1. **Sélection des retours utilisateur (Feedback Identification)** : Déterminer un ensemble $P_t \subseteq \mathcal{A}$ pour lequel des retours seront attendus. Pour les bandits en cascade [Hir+19; Kve+15], cela définit un critère d'arrêt (ex. sélection d'un élément préféré a_s dans S_t). D'autres approches considèrent $P_t = \mathcal{A}$ ou $P_t = S_t$.
2. **Acquisition du retour utilisateur (Feedback Retrieval)** : Construire un vecteur initial F_t . Pour les bandits en cascades, cela se fait via des distances relatives de chaque bras a à a_s . D'autres méthodes sollicitent explicitement un retour pour chaque bras de P_t ou une opinion globale sur S_t .
3. **Calcul de la récompense (Reward Computing)** : Calculer R_t (scalaire ou vecteur) utilisé pour mettre à jour la politique de l'agent. R_t peut être F_t ou le résultat d'un traitement sur F_t (pondération, produit scalaire, pseudo-récompenses [Gup+21]).

Ces processus peuvent être ajustés pour maximiser les performances des algorithmes selon la tâche à résoudre.

Sélection des retours utilisateur : apprentissage actif L'approche classique pour les retours utilisateur dans une recommandation S_t est de sélectionner les $\psi \leq k$ éléments avec la plus haute espérance de récompense [Lou16]. Cependant, des stratégies d'apprentissage actif peuvent être utilisées pour solliciter des retours sur des éléments moins explorés, améliorant ainsi l'apprentissage à long terme [ERR16 ; Ela+13].

Calcul de la récompense : valorisation des retours Les récompenses R_t sont essentielles pour guider l'apprentissage des algorithmes de bandits [SB98]. Récemment, GUPTA et al. [Gup+21] ont proposé des "pseudo-récompenses" pour les bras non recommandés, fondées sur des corrélations préalables, méthode pouvant être étendue aux recommandations multiples.

Ces approches montrent que la conception des fonctions de récompense est cruciale pour optimiser les performances des algorithmes de bandits, tant en précision qu'en temps de convergence.

1.2 Positionnement des contributions

Pour mieux appréhender le cœur de nos contributions concernant les systèmes de recommandation bi-objectifs basés sur les modèles de *MAB* et de *CMAB*, nous avons rappelé dans la section précédente leurs principales définitions, que ce soit dans le cadre de tirages simples ou multiples.

Dans cette section, nous évoquons nos contributions relatives aux systèmes de recommandation bi-objectifs basés sur les *MAB/CMAB* avec tirages simples ou multiples, que nous avons implémentés sous la forme d'une suite de méthodes de type portfolio d'algorithmes de recommandation, appelée *Gorthaur* [Gut19 ; Gut+19b ; Gut+21 ; Let+24].

Les systèmes de recommandation à tirage simple ou multiple reposant sur les bandits-manchots constituent un domaine dans lequel les approches multi-objectifs trouvent leur place de manière pertinente. Cependant, ces applications soulèvent des problématiques similaires à celles rencontrées dans des travaux plus généralistes, telles que l'utilisation ou non de techniques de scalarisation, la hiérarchisation des objectifs, ou encore la détermination d'un front de Pareto adapté aux besoins d'optimisation. Par ailleurs, tout comme les systèmes de recommandation mono-objectifs, ces approches évoluant dans des environnements dynamiques peuvent être confrontées à des problématiques telles que la parcimonie des données (sparsity), la dérive conceptuelle (concept-drift) ou le changement de concept (concept-shift). Face à ces enjeux, nous avons envisagé deux stratégies reposant sur différents types d'approches nouvelles pour palier les biais induits par les données, leur dérive ou leur instabilité, ainsi que leur parcimonie ou leur forte dimensionnalité :

1. **Développer des modèles robustes capables de s'adapter à ces problèmes intrinsèques sans modifier les données elles-mêmes.** Cela a impliqué la mise en œuvre

de méthodes telles que *Gorthaur-EXP3* [Gut+21], *Multiple Play Gorthaur-TS* ou encore *Budgeted Gorthaur-EXP3* [Let+24], qui exploitent des algorithmes de bandits pour sélectionner dynamiquement le meilleur algorithme de recommandation dans un portefeuille, en réponse aux variations de l'environnement et aux besoins de précision et de diversité. Ces approches sont particulièrement résistantes aux phénomènes de concept-drift et de concept-shift, ainsi qu'à la non-stationnarité. Elles permettent de gérer efficacement des données parcimonieuses tout en maintenant une précision optimale. Ces travaux de recherches sur les algorithmes de recommandation bi-objectifs adaptés aux environnements dynamiques et données incertaines ont fait l'objet de deux contributions :

- En 2021, dans le journal *Information Sciences* [Gut+21] (Elsevier) ;
- En 2024, dans le journal *Expert Systems with Applications* [Let+24] (Elsevier)

2. **Palier la parcimonie des données pour les rendre plus exploitables par les modèles d'apprentissage.** Cela a impliqué l'application de stratégies telles que l'approche *Partial Bandit with Semi-Bandit (P-BSB)* [Let+20b] pour optimiser l'utilisation des retours d'utilisateurs limités, ainsi que la méthode *Bandit Under Semi-Bandit Conditions (BUSBC)* [Let+22], qui améliore la précision des algorithmes de bandits-manchots combinatoires en adaptant le calcul des récompenses à des conditions de retours utilisateur partiels. Ces approches permettent de tirer parti de données parcimonieuses tout en maintenant une performance robuste face à l'incertitude et à la variabilité des données. Ces travaux de recherches ont fait l'objet de deux articles en conférence :

- En 2020, dans la conférence *IEEE ICTAI* [Let+20b] (International Conference on Tools with Artificial Intelligence) ;
- En 2022, dans la conférence *FLAIRS* [Let+22] (The Florida Artificial Intelligence Research Society)

Bien que ces dernières contributions concernant la prise en compte de retours partiels soient complémentaires et utiles pour nos approches de systèmes de recommandation bi-objectifs, elles restent à la marge de la thématique principale de ce chapitre de la Partie I portant sur l'optimisation multicritère. Nous ne les traiterons donc pas directement dans ce chapitre et renvoyons pour plus d'informations à l'Annexe A.1 où elles sont décrites.

Les contributions que nous mettrons en lumière par la suite se focalisent donc sur notre suite de méthodes de type portfolio d'algorithmes, dénommée *Gorthaur*. Elles correspondent à de la méta-sélection d'algorithmes d'apprentissage pour la recommandation bi-objectif. C'est pourquoi, dans la section suivante, nous rappelons un état de l'art sur la méta-sélection d'algorithmes d'apprentissage.

1.3 Méta-sélection d'algorithmes d'apprentissage

Cette section dresse un rapide état de l'art sur la méta-sélection d'algorithmes d'apprentissage. Pour ce faire, elle repose sur des éléments présentés dans [Gut+21 ; Let23 ; Let+24].

1.3.1 Motivation

Les programmes informatiques peuvent résoudre divers problèmes, tels que la classification, la régression, la planification, et l'optimisation. Toutefois, il est bien établi qu'aucun algorithme générique ne surpasse tous les autres pour l'ensemble des problèmes et critères associés, comme l'indiquent le *No Free Lunch Theorem (NFLT)* et la loi de conservation de la performance moyenne [HP01 ; Ker+19].

Un algorithme optimal pour un problème donné est celui qui obtient les meilleures performances moyennes pour une instance de ce problème. La définition de l'algorithme optimal doit donc être adaptée selon le cas, en fonction :

- des critères à optimiser, tels que la précision, le temps de calcul, la diversité ou la fiabilité ;
- du type d'optimisation ciblée, qu'il s'agisse d'améliorer la performance moyenne ou la performance du pire cas (minimax) ;
- des contraintes spécifiques de l'application.

L'efficacité d'un algorithme peut varier d'une instance de problème à une autre, selon les critères mentionnés. Ainsi, déterminer à l'avance la méthode optimale pour une nouvelle instance est complexe. De plus, les progrès de la recherche ont multiplié les méthodes disponibles, rendant la sélection de l'approche la plus appropriée encore plus difficile. C'est à ce titre que les approches de méta-apprentissage prennent tout leur sens.

Le méta-apprentissage inclut la sélection d'algorithmes, visant à améliorer les performances de résolution de problèmes en étudiant les méthodes utilisées. Nous en décrivons le concept général à la sous-section suivante.

1.3.2 Concept de la sélection automatique d'algorithmes

Le problème de la sélection d'un algorithme efficace a été introduit par Rice [Ric+76]. Son idée est de *mapper* chaque couple problème-algorithme en fonction des performances de l'algorithme sur un problème donné (instance). Ainsi, il définit le problème comme suit (voir Figure 1.2) : Soit $d \in \mathcal{D}$ une instance particulière du problème et $h(d) \in \mathcal{H}$ la fonction d'extraction de données correspondant aux méta-données caractérisant les instances de \mathcal{D} , soit ω l'ensemble des critères à optimiser et soit \mathcal{B} l'ensemble des algorithmes considérés pour résoudre le problème, et enfin soit z l'ensemble des performances en regard des critères considérés. Le problème de sélection d'algorithmes peut être défini comme la recherche d'une fonction associative $S(h(d), \omega)$ pour une instance $d \in \mathcal{D}$ d'un problème spécifique, caractérisée par ses dimensions $h(d) \in \mathcal{H}$. L'objectif est de déterminer l'algorithme $b \in \mathcal{B}$, \mathcal{B} étant un ensemble

d'algorithmes, qui maximise la performance $\|z\| = g(f(b, d), \omega)$, selon l'ensemble des critères d'optimisation $\omega \in \mathbb{R}^n$.

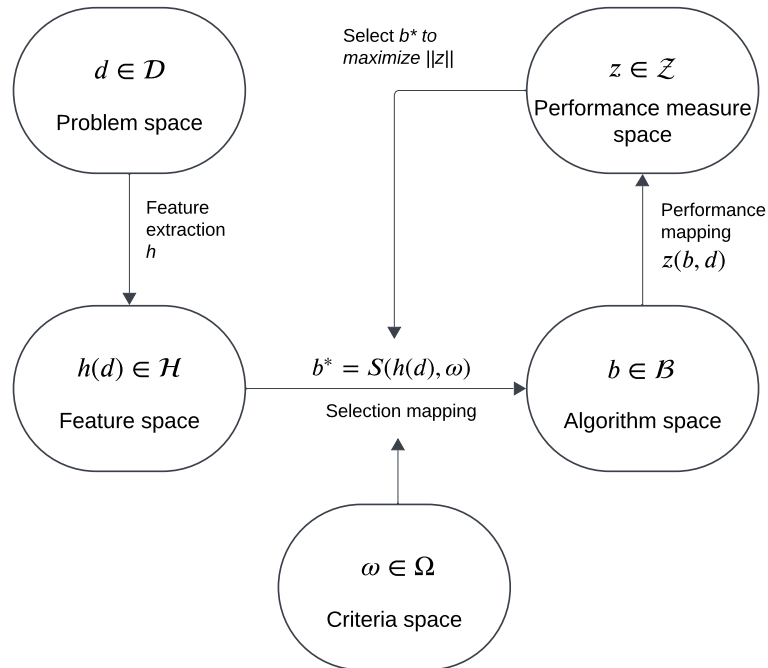


FIGURE 1.2 – Schéma du modèle du problème de sélection d'algorithmes de Rice [Ric+76] appliquée au multicritère

1.3.3 Cas de la sélection automatique d'algorithmes pour la recommandation

Pour maximiser les performances des systèmes de recommandation en ligne, une solution, de type approche portfolio, pourrait consister à choisir le meilleur algorithme parmi ceux de l'état de l'art. Cependant, sans connaissance préalable, il est impossible de sélectionner l'algorithme optimal pour tous les cas. Il est donc nécessaire d'utiliser un mécanisme de sélection dynamique pour choisir le candidat le mieux adapté à une application ou un jeu de données spécifique.

Dans le contexte des problèmes de bandits-manchots (*Multi-Armed Bandit - MAB*) avec une évaluation en ligne, diverses approches ont été développées, et la méta-sélection d'algorithmes (approche portfolio) consiste souvent à choisir dynamiquement un algorithme dans un portefeuille [GS06 ; LF18 ; Smi09].

Une fois le choix des algorithmes à inclure dans le portefeuille effectué, un mécanisme de sélection approprié doit ensuite être défini. À titre d'exemple, nous rappelons ci-dessous deux stratégies populaires de sélection qui ont inspiré nos approches.

Dans [AF17], une nouvelle approche de méta-sélection est proposée, basée sur la sélection d'« experts en apprentissage » à l'aide de l'algorithme *EXP3*. Chaque expert en apprentissage

est défini comme un algorithme de bandit-manchoth contextuel. Un agent doit choisir un expert, lequel sélectionne une action à chaque itération après avoir observé un vecteur de contexte. Chaque expert cherche à minimiser son erreur d'estimation, tandis que l'agent évalue la performance des experts.

Les auteurs démontrent que l'exploration est assurée par des algorithmes efficaces, et que le biais des experts, tel que la sensibilité au contexte, est réduit en sélectionnant le meilleur d'entre eux.

Deux algorithmes ont été proposés pour gérer la méta-sélection : *Learn, Then Explore and Exploit (LTEE)* et *Learn, Explore and Exploit (LEE)*. *LTEE* est un algorithme opérant en deux étapes : 1) une phase d'apprentissage, où chaque expert est évalué ; 2) une phase où un expert optimal est sélectionné en éliminant successivement les moins performants. Bien que cet algorithme soit théoriquement performant, il peut être affecté par la complexité de l'expert le moins efficace. Pour améliorer les résultats, l'algorithme *LEE* quant à lui gère simultanément l'apprentissage, l'exploration et l'exploitation, réduisant ainsi la complexité et améliorant les performances.

Dans la section suivante, nous détaillons le fonctionnement des nos méthodes *Gorthaur*.

1.4 Méthodes *Gorthaur* : Méta-apprentissage pour la recommandation

De fin 2019 à juillet 2024, nous avons développé une suite de méthodes nommée *Gorthaur*[Gut+19b ; Gut+21 ; Let+24], un acronyme pour :

Generic-ORienTed Heuristic Algorithm for User Recommendation.

Ces méthodes reposent sur un portefeuille dynamique d'algorithmes pour sélectionner de manière adaptative l'algorithme de recommandation le plus approprié face aux enjeux spécifiques de l'environnement (tels que le concept-shift, la non-stationnarité, ou la parcimonie), tout en optimisant les critères de précision et de diversité des recommandations via une méthode de détermination d'un front de Pareto utilisant un *Compass* [MS08]. La méthode *Compass* (2008) est utilisée dans les algorithmes génétiques comme une approche adaptative qui vise à équilibrer l'exploration et l'exploitation de l'espace de recherche. Elle repose sur une abstraction des paramètres des opérateurs, en évaluant leurs effets sur la diversité et la qualité moyenne de la population. Le *Compass* ajuste dynamiquement les probabilités d'application des opérateurs pour guider la recherche en fonction d'un paramètre unique, Θ , qui contrôle cet équilibre, facilitant ainsi l'évasion des optima locaux et l'amélioration des solutions. Outre avoir été utilisée dans la suite de méthodes *Gorthaur*, cette méthode est encore régulièrement utilisée [FT23 ; Li+24 ; Pei+23 ; TFF24] ou identifiée dans les revues de l'état de l'art récentes du domaine [Kar+22].

Le développement de la suite de méthodes *Gorthaur* repose sur deux constats majeurs :
1) les systèmes de recommandation doivent être robustes aux variations de l'environnement

d'apprentissage; 2) ils doivent non seulement viser une haute précision, mais aussi assurer une diversité suffisante dans les recommandations [Gut19]. Bien que les algorithmes de bandits-manchots possèdent de solides fondements théoriques, leurs performances peuvent varier selon les contextes d'application ou les jeux de données utilisés. Par conséquent, il est crucial de mener une évaluation préliminaire hors-ligne sur des critères comme la précision et, le cas échéant, la diversité, avant de sélectionner un algorithme. En réalisant la sélection d'algorithmes de bandits-manchots pour la recommandation de manière dynamique au cours du temps, la suite de méthodes *Gorthaur* permet de s'émanciper de cette phase coûteuse, voire souvent impossible à mener en raison de la nature interactive, évolutive et dynamique des systèmes de recommandation.

Cette section vise à décrire la suite de méthodes *Gorthaur* que nous avons développée. Pour ce faire, elle rappelle en préambule les travaux préliminaires initiés dans [Gut19 ; Gut+19b] et décrit ensuite l'ensemble des éléments qui en découlent présentés dans [Gut+21 ; Let23 ; Let+24]. De même, cette section généralise les méthodes *Gorthaur* utilisant des *MAB* comme sélecteur, en définissant le problème formellement. Enfin, pour chaque méthode *Gorthaur-MAB* clé, elle revient sur la preuve théorique en en proposant une schématique (Sketch) à partir de celle originale.

1.4.1 Cœur des méthodes *Gorthaur* : le Compass

Gorthaur

Un algo pour les gouverner tous,
Un algo pour les trouver,
Un algo pour les amener tous,
Et dans le multicritère les lier.
Dérivé de [Led72]...

Selon la littérature sur les systèmes de recommandation, l'évaluation des algorithmes de bandit-manchot est le plus souvent basée sur les récompenses cumulées (ou le regret) qu'ils obtiennent. Cela signifie que de tels systèmes de recommandation visent à maximiser la précision globale, ce qui, dans ce cas d'application spécifique, peut se faire au détriment d'une faible diversité. Le principe des méthodes *Gorthaur* [Gut+21] repose sur l'utilisation d'un portefeuille d'algorithmes de *MAB* et de *CMAB* pour la recommandation et utilise une stratégie spécifique pour les sélectionner (p.ex., roulette, *MAB*, *CMAB*). L'objectif des méthodes *Gorthaur* est de maximiser le compromis entre la précision globale et la diversité selon une cible donnée. Le principal objectif des méthodes *Gorthaur*, inspiré par la méthode *Compass* [MS08], est donc de tirer parti de l'avantage de chaque algorithme du portefeuille sur ces deux critères.

Dans le cadre de nos travaux, le problème d'optimisation bi-objectifs peut être formulé de la manière suivante :

$$\max(acc(t), div(t)) \text{ s.c. } t \in [1, T]$$

où t représente l'itération (pas de temps) et T l'horizon.

Comme présenté dans [Gut+21] et illustré dans la Figure 1.3, soit ΔAcc la variation de la précision globale sur l'axe vertical, et ΔDiv la variation de la diversité sur l'axe horizontal. Au point de départ t_0 , nous exprimons un vecteur de référence \vec{c} défini par l'angle $\Theta \in [0; \frac{\pi}{2}]$ formé par \vec{c} avec l'axe horizontal (ΔDiv). Ce vecteur de référence \vec{c} exprime le compromis requis entre la précision globale et la diversité. En fonction de la valeur de Θ que nous définissons, nous pouvons choisir de privilégier la précision globale, la diversité, ou de calculer un équilibre entre les deux critères.

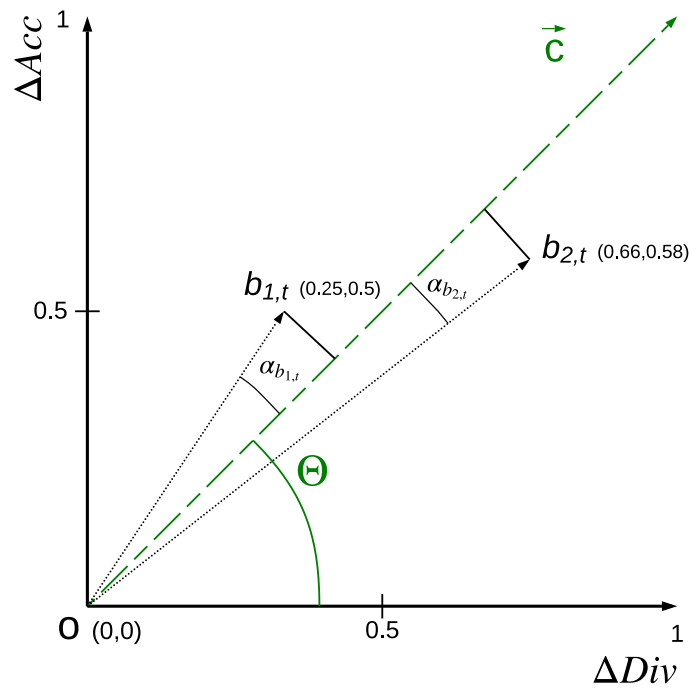


FIGURE 1.3 – Projections de *Gorthaur* sur le vecteur de référence \vec{c} pour calculer la pertinence [Gut+21]

1.4.2 Méthodes *Gorthaur* à tirage simple

Avant de proposer une définition formelle générale des méthodes *Gorthaur*, nous en rappelons le fonctionnement ci-dessous.

1.4.2.1 Fonctionnement des méthodes *Gorthaur* à tirage simple

Soit \mathcal{B} un ensemble d'algorithmes de recommandation, de *MAB* et de *CMAB*, ayant tous le même nombre fixe d'éléments à recommander (bras du bandit). À chaque itération t , *Gorthaur* sélectionne un algorithme $b \in \mathcal{B}$ qui choisit lui-même un élément a à recommander selon sa propre stratégie. Ensuite, *Gorthaur* calcule la précision globale $acc(b, t)$ de l'algorithme b à l'itération t et sa diversité $div(b, t)$.

$acc(b, t)$, la précision globale moyenne est définie par :

$$acc(b, t) = \frac{g(b, t)}{t_b} \quad (1.1)$$

où $g(b, t) = \sum_{t=1}^T R(b, t)$ est la somme des récompenses obtenues par l'algorithme b à l'itération t , $\forall t \in [1; T]$, et t_b est le nombre de fois où l'algorithme b a été sélectionné par *Gorthaur* depuis l'itération initiale t_0 .

$div(b, t)$ est la diversité obtenue par l'algorithme b à l'itération t et est définie comme suit. En considérant $N_{b,t} = \{n_{a_1}(b, t), \dots, n_{a_m}(b, t)\}$ où $n_{a_j}(b, t)$ est le nombre total de fois où le bras a_j a été recommandé par l'algorithme b à l'instant t , nous définissons formellement $c_v(N_{b,t}) = \frac{\sigma(N_{b,t})}{\bar{N}_{b,t}}$ où $\bar{N}_{b,t}$ est le nombre moyen de fois où un bras a été sélectionné par l'algorithme b à l'itération t , et $\sigma(N_{b,t})$ est son écart type. Ensuite, la diversité résultante pour chaque algorithme b à l'itération t est :

$$div(b, t) = 1 - \frac{c_v(N_{b,t})}{\sqrt{k}} \quad (1.2)$$

comme défini par [Gut+19a].

La capacité d'un algorithme à fournir des recommandations à la fois précises et diversifiées est alors exprimée par un vecteur $o_{b,t}$ tel que :

$$o_{b,t} = (div(b, t), acc(b, t))$$

Les valeurs mesurées de $div(b, t)$ et $acc(b, t)$ sont alors normalisées comme suit :

$$div^{norm}(b, t) = \frac{div(b, t)}{\max\{div(b, t)\}}$$

et

$$acc^{norm}(b, t) = \frac{acc(b, t)}{\max\{acc(b, t)\}}$$

Cela permet de calculer un vecteur normalisé :

$$o_{b,t}^{norm} = (div^{norm}(b, t), acc^{norm}(b, t)) \quad (1.3)$$

La méthode *Gorthaur* poursuit ensuite par l'étape de sélection d'algorithmes en utilisant au choix de l'implémentation :

- soit une méthode de roulette proportionnelle [Gol89; LL12; Thi05] (*Gorthaur-Wheel*), qui est notre proposition originale de fin 2019 [Gut+19b];
- soit une méthode reposant sur un algorithme de *MAB* (*Gorthaur-MAB*), p.ex., comme *EXP3* (*Exponential-weight algorithm for Exploration and Exploitation*) [Aue+95; Aue+02] qui est notre proposition de 2021 [Gut+21], ou encore comme *Thompson Sampling* (*TS*) [AG12] qui est l'une de nos propositions de 2024 [Let+24] (avec une spécificité

supplémentaire puisqu'elle est couplée à un algorithme Multiple-play dans le cadre de recommandations de multiples éléments, voir la sous-section 1.4.4).

Ces deux types de méthode de sélection reposent sur la fitness $f_{b,t}$ de chaque algorithme. En effet, la fitness $f_{b,t}$ est obtenue en utilisant la projection du vecteur normalisé $o_{b,t}^{norm}$ sur le vecteur de référence \vec{c} (voir Figure 1.3). Il est donc possible de calculer la fitness $f_{b,t}$ pour chaque algorithme $b \in \mathcal{B}$ à l'itération t comme suit :

$$f_{b,t} = |o_{b,t}^{norm}| \cos \alpha_{b,t} - \min_b \{|o_{b,t}^{norm}| \cos \alpha_{b,t}\} \quad (1.4)$$

Ensuite, à chaque itération t , *Gorthaur* choisit un algorithme $b \in \mathcal{B}$ selon la méthode de sélection employée :

- Pour la méthode d'origine *Gorthaur-Wheel* [Gut+19b] avec une probabilité $p_{b,t}$ définie comme suit :

$$p_{b,t} = \frac{f_{b,t} + \zeta}{\sum_{i=1}^{|\mathcal{B}|} f_{b_i,t} + \zeta} \quad (1.5)$$

où $\zeta = 2^{-1074}$ est une constante qui évite à la fois de diviser par 0 et garantit qu'à chaque itération, chaque algorithme a une probabilité minimale d'être sélectionné par la roulette.

- Pour *Gorthaur-MAB* [Gut+21 ; Let+24] le choix de l'algorithme b est effectuée en fonction de la politique de sélection de l'algorithme de *MAB*. Par exemple pour *Gorthaur-EXP3*, avec une probabilité $p_{b,t}$ définie comme suit :

$$p_{b,t} = (1 - \eta) \frac{w_{b,t}}{\sum_{i=1}^{|\mathcal{B}|} w_{b_i,t}} + \frac{\eta}{|\mathcal{B}|} \quad (1.6)$$

où $\eta \in [0, 1]$ représente le paramètre d'exploration et $w_{b,t}$ le poids pour l'algorithme b à l'itération t et mis à jour selon l'Équation 1.7.

La sélection du bras/élément $a \in A$ est ensuite traitée selon la stratégie de l'algorithme b choisi par *Gorthaur*. Enfin, la récompense R_t est observée et les paramètres de l'algorithme b du portefeuille sont mis à jour selon leur propre politique.

Enfin, *Gorthaur* observe la récompense obtenue R_t et procède à la mise à jour de la précision, de la diversité et de la fitness $f_{b,t}$ de l'algorithme b_t (Voir Équations 1.1 à 1.4).

Une étape supplémentaire de mise à jour des récompenses de l'algorithme de *MAB* qui le meta-sélecteur dans le cadre de *Gorthaur-MAB* est nécessaire. Par exemple pour *Gorthaur-EXP3* : pour l'algorithme choisi b_t , *Gorthaur-EXP3* calcule $\hat{f}_{b,t} = \frac{f_{b,t}}{p_{b,t}}$ et pour les autres algorithmes il applique $\hat{f}_{b,t} = 0$ (c'est-à-dire, pour les algorithmes qui n'ont pas été sélectionnés à l'itération t). Enfin, *Gorthaur-EXP3* met à jour ses poids comme suit :

$$\forall b \in |\mathcal{B}|, w_{b,t+1} = w_{b,t} \exp \left(\frac{\eta}{|\mathcal{B}|} \hat{f}_{b,t} \right) \quad (1.7)$$

1.4.3 Problème de bandit-manchot de sélection d'algorithmes multicritère : Définitions

Dans nos travaux précédents, nous avons tout d'abord (fin 2019) proposée une nouvelle approche de sélection d'algorithmes de bandits-manchots pour la recommandation à l'aide d'une roulette proportionnelle [Gut+19b]. Nous avons ensuite (2021) amélioré cette méthode tant sur le plan empirique que théorique en modifiant son mécanisme de sélection roulette par l'algorithme de *MAB EXP3* [Gut+21].

Enfin, pour conclure sur les méthodes *Gorthaur* à tirage simple, nous avons proposé de généraliser cette sélection pour permettre l'utilisation de tout algorithme de bandit-manchot (*Multi-Armed Bandit - MAB*) comme méta-algorithme [Let+24].

Le principe de fonctionnement des méthodes *Gorthaur* à tirage simple ayant été rappelé à la sous-section précédente, nous établissons la formalisation du problème au travers des Définitions 10, 11 et 12.

Le problème de la sélection d'algorithmes peut être formulé comme un problème de bandit-manchot (*Multi-Armed Bandit - MAB*) comme suit :

Définition 10. Le problème de bandit-manchot (*Multi-Armed Bandit - MAB*) pour la sélection d'algorithmes dans les problèmes de recommandation multicritères. Soit \mathcal{B} l'ensemble des algorithmes d'un portefeuille et $b \in \mathcal{B}$ un algorithme spécifique du portefeuille où les algorithmes de \mathcal{B} sont associés à $|\mathcal{B}|$ bras indépendants à tirer à chaque instant $t = \{1, 2, 3, \dots\}$ de ce problème de bandit stochastique à $|\mathcal{B}|$ bras.

Lorsqu'un algorithme b_i est sélectionné, il reçoit une récompense $r_{b_i,t} \in [0, 1]$ obtenue à partir d'une fonction d'agrégation normalisée $f_{b_i}^{norm} \in [0, 1]$ qui correspond à la performance globale atteinte par b_i selon les critères à optimiser. La récompense est observée immédiatement après que $b_i \in \mathcal{B}$ est sélectionné et appliqué.

L'objectif de l'agent, appelé méta-algorithme, est de minimiser la perte encourue en tirant des bras sous-optimaux, c'est-à-dire en appliquant des algorithmes sous-optimaux pour résoudre l'instance du problème.

Définition 11. Algorithme *Gorthaur-MAB* pour la sélection d'algorithmes dans les problèmes de recommandation. Dans un problème de recommandation typique, où la méthode de résolution optimale est inconnue, un agent, appelé méta-algorithme, vise à résoudre le problème de sélection d'algorithmes. À chaque itération t , le méta-algorithme détermine un algorithme $b_i \in \mathcal{B}$, $\mathcal{B} = \{b_1, \dots, b_{|\mathcal{B}|}\}$, à utiliser, basé sur les résultats des $t - 1$ observations précédentes $\{(b_{i,1}, r_{b_i,1}), \dots, (b_{i,t-1}, r_{b_i,t-1})\}$.

Un algorithme *Gorthaur-MAB* inclut un algorithme de *MAB* comme méta-algorithme de sélection et calcule $f_{b_i}^{norm} \in [0, 1]$ comme étant la valeur de fitness entre les performances atteintes de l'algorithme b_i et le compromis désiré entre les critères. À la réception de la valeur de fitness de l'algorithme, l'agent met à jour le vecteur des récompenses \vec{r}_t .

L'objectif de l'agent est de maximiser la récompense totale attendue $\sum_{t=1}^T \mathbb{E}_{\vec{r}_t \sim D[r_{b,t}]}$.

Définition 12. Regret dans la sélection d’algorithmes Gorthaur-MAB. Soit D_r la distribution de l’espérance de récompense des algorithmes à sélectionner. Par conséquent, $D_r = (\mu_{b_1}, \dots, \mu_{b_{|\mathcal{B}|}}) \in [0, 1]^{|\mathcal{B}|}$, où $\mu_{b_i} \in [0, 1]$ est la récompense attendue lors de la sélection de l’algorithme b_i , $i \in [1, |\mathcal{B}|] \cap \mathbb{N}$.

L’objectif est donc de maximiser la récompense totale attendue $\mathbb{E}[\sum_{t=1}^T \mu_{b_i(t)}]$ à l’horizon T , où $b_i(t)$ est l’algorithme sélectionné et appliqué à l’itération t . Néanmoins, la littérature sur les problèmes de bandit-manchot utilise la mesure équivalente au regret total attendu, qui est plus pratique. Elle reflète essentiellement la quantité que l’on perd en ne sélectionnant pas l’algorithme optimal à chaque itération t . Soit $\mu^* = \max_{b_i \in \mathcal{B}} \mu_{b_i}$ et $\Delta_{b_i} = \mu^* - \mu_{b_i}$. Soit $n_{b_i}(t)$ le nombre de fois où l’algorithme b_i a été joué jusqu’à l’itération $t - 1$. Alors le regret total attendu d’un Gorthaur-MAB à l’horizon T est :

$$\mathbb{E}[\rho_{\text{Gorthaur},f}(T)] = \mathbb{E}\left[\sum_{t=1}^T (\mu^* - \mu_{b_i(t)})\right] = \sum_i \Delta_{b_i} \cdot \mathbb{E}[n_{b_i}(T)]$$

1.4.3.1 Le portefeuille possible de Gorthaur

En suivant les différents critères des algorithmes en termes de précision, personnalisation, diversité et applicabilité aux applications en temps réel, nous proposons un portefeuille de Gorthaur composé des 6 algorithmes de bandit-manchot suivants, à la fois contextuels et non contextuels :

- **Algorithmes de bandit-manchot (Multi-Armed Bandit - MAB) :** UCB1 [Aue+02], Thompson Sampling (TS) [AG12], et EXP3 [Aue+02]
- **Algorithmes de bandit-manchot Contextuels (Contextual Multi-Armed Bandit CMAB) :** LinUCB [Li+10], SW-LinUCB [Gut+19a], et Linear Thompson Sampling (LinTS) [AG13b].

Cet ensemble spécifique d’algorithmes a été sélectionné en fonction des besoins particuliers. Néanmoins, il est possible d’utiliser d’autres types d’algorithmes de recommandation dans le portefeuille des méthodes Gorthaur, par exemple, les algorithmes de filtrage collaboratif [RRS21], les algorithmes de la chaîne de Markov de Monte-Carlo [AEK00].

1.4.3.2 Les promesses des méthodes Gorthaur

Notre méthode d’origine Gorthaur-Wheel [Gut+19b] sélectionne les algorithmes (parmi son portefeuille) qui correspondent le mieux aux ensembles de données ou aux cas qu’il rencontre. C’est un avantage majeur lorsqu’on traite en ligne avec des applications du monde réel. Les performances d’une telle approche dépendent donc à la fois de 1) des algorithmes du portefeuille et 2) de la stratégie utilisée pour sélectionner un algorithme du portefeuille.

Les premières expériences avec Gorthaur-Wheel rapportées dans [Gut+19b] visant à valider l’approche de portefeuille mise en œuvre dans Gorthaur-Wheel ont montré que, dans le cas de la recommandation, la méthode a réussi à faire face à différents types d’ensembles de données. Bien sûr, Gorthaur-Wheel a choisi des algorithmes de CMAB lorsqu’il a été confronté

à des configurations contextuelles et des algorithmes de *MAB* pour des problèmes non contextuels. De plus, avec les ensembles de données sur lesquels il a été expérimenté, *Gorthaur-Wheel* a pu choisir les algorithmes les plus efficaces pour atteindre le compromis requis de précision/diversité.

Ces résultats positifs nous ont ainsi conduits à : 1) étendre et perfectionner la méthode *Gorthaur*, d'abord pour le tirage simple en 2021 avec un algorithme de sélection *EXP3* (*Gorthaur-EXP3* [Gut+21]); 2) à généraliser ensuite cette approche pour tout algorithme de *MAB* [Let+24] en 2024 tout en proposant de décliner cette approche à du tirage multiple [Let+24] (voir Sous-section 1.4.4).

Pour illustrer le fonctionnement de *Gorthaur-MAB* à tirage simple, nous décrivons notre contribution [Gut+21] ci-dessous.

1.4.3.3 Méthode *Gorthaur-EXP3*

À la suite des résultats prometteurs obtenus par *Gorthaur-Wheel*, nous avons créé et expérimenté *Gorthaur-EXP3*. Comme expliqué précédemment, la transition de *Gorthaur-Wheel* vers *Gorthaur-EXP3* repose sur le passage d'une sélection basée sur la roulette, qui évalue chaque algorithme en fonction de sa pertinence, à une approche où la sélection des algorithmes du portefeuille est formulée comme un problème de bandits-manchots (*MAB*). Pour résoudre ce problème, nous utilisons l'algorithme *EXP3* spécifiquement conçu pour les problèmes de *MAB*.

Algorithme *Gorthaur-EXP3*. Nous avons précédemment défini les différentes étapes de fonctionnement des méthodes *Gorthaur* à tirage simple. Pour plus de clarté, nous donnons ici le pseudo code de *Gorthaur-EXP3*. Ainsi, un système de recommandation utilisant *Gorthaur-EXP3* fonctionne comme montré dans l'algorithme 1.

Notons que l'angle du vecteur Θ a été fixé ici à 45° . Cependant, celui-ci est paramétrable (voir Figure 1.4) ou peut être laissé en mode de calcul dynamique, permettant à l'algorithme de déterminer lui-même, au fil des itérations, le front de Pareto en optimisant les compromis entre précision et diversité.

Regret de *Gorthaur-EXP3*. À la différence d'une sélection roulette proportionnelle, l'utilisation d'*EXP3* nous permet de reposer sur des garanties théoriques.

Comme l'a prouvé [LF18], lorsqu'on utilise une approche de portefeuille avec des algorithmes d'apprentissage par renforcement, le regret ne serait pas pire que celui du pire algorithme. De plus, selon le No Free Lunch Theorem, il ne serait pas non plus meilleur que le meilleur algorithme du portefeuille [HP02].

Cependant, comme la fonction de récompense repose entièrement sur le calcul de la pertinence de *Gorthaur-EXP3*, nous pouvons nous assurer que *Gorthaur-EXP3* présente une convergence similaire à l'algorithme qui correspond le mieux au compromis précision/diversité

cible. De plus, le regret final de *Gorthaur-EXP3* converge vers celui de son meilleur algorithme proportionnellement à sa probabilité de sélection.

Nous nous attendons à ce que *Gorthaur-EXP3* soit plus sélectif que sa version précédente avec sélection par roulette selon le Théorème de convergence suivant :

Théorème 1. Convergence de Gorthaur-EXP3 sur le compromis précision/diversité. Soit B l'ensemble des algorithmes du portefeuille de *Gorthaur-EXP3* et $b_i \in B$ un algorithme spécifique de ce portefeuille où $i \in [1; |B|] \subseteq \mathbb{N}$. Soit $f_{b_i}(t)$ la pertinence de l'algorithme $b_i \in B$ à l'itération $t \in [1; T]$ où T est l'horizon fini. Soit $\eta \in [0; 1]$ le facteur d'exploration à initialiser au point de départ $t_0 = 0$.

Algorithme 1 : - Algorithme *Gorthaur-EXP3* pour la recommandation

Require: Liste des utilisateurs $u \in U$ et de leur contexte $x \in X$ (si le contexte est disponible).

Liste des k éléments à recommander associés aux bras $a \in A$. Le portefeuille d'algorithmes $b \in B$. Angle $\Theta = 45^\circ$. $\eta \in [0; 1]$. $w_{b_{i,1}} = 1$ pour $i = 1, \dots, |B|$.

1: **for** $t = 1$ à T **do**

2: Sélectionner aléatoirement un utilisateur $u_t \in U$ et son contexte $x_t \in X$

3: **for all** $b_i \in B$ **do**

4: Calculer $p_{b_{i,t}} = (1 - \eta) \frac{w_{b_{i,t}}}{\sum_{j=1}^{|B|} w_{b_{j,t}}} + \frac{\eta}{|B|}$ comme défini dans l'Équation (1.6)

5: **end for**

6: Effectuer la sélection de l'algorithme b_t aléatoirement selon les probabilités

$p_{b_{1,t}}, \dots, p_{b_{|B|,t}}$, puis exécuter l'algorithme b_t

7: Sélectionner l'élément $a \in A$ selon la stratégie de l'algorithme précédemment choisi b_t et recommander cet élément à l'utilisateur u_t

8: Observer la récompense obtenue R_t

9: Mettre à jour les paramètres de l'algorithme précédemment choisi b_t selon sa stratégie de traitement des récompenses

10: Mettre à jour $acc(b_t, t)$ comme défini dans l'Équation (1.1)

11: Mettre à jour $div(b_t, t)$ comme défini dans l'Équation (1.2)

12: Mettre à jour $\sigma_{b_{i,t}}^{norm}$ comme défini dans l'Équation (1.3)

13: Mettre à jour $f_{b_{i,t}}$ comme défini dans l'Équation (1.4)

14: **for** $j = 1$ à $|B|$ **do**

15:
$$\hat{f}_{b_{j,t}} = \begin{cases} \frac{f_{b_{j,t}}}{p_{b_{j,t}}} & \text{si } b_j = b_t \\ 0 & \text{sinon} \end{cases}$$

16: Mettre à jour $w_{b_{j,t+1}} = w_{b_{j,t}} \exp\left(\eta \frac{\hat{f}_{b_{j,t}}}{|B|}\right)$

17: **end for**

18: **end for**

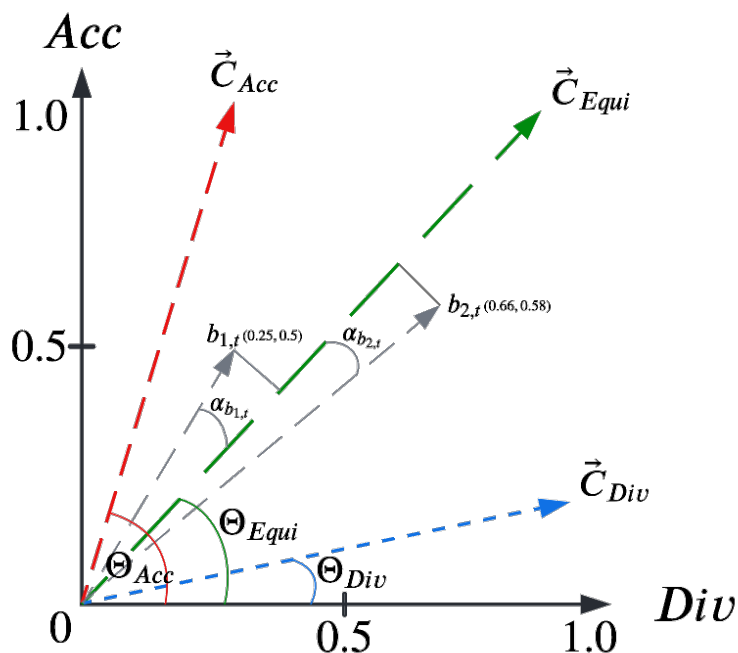


FIGURE 1.4 – Différents compromis entre précision et diversité du Compass de Gorthaur [Gut19 ; Gut+21 ; Let+24]

Alors, selon la preuve générale de EXP3 présentée dans [Aue+02], nous pouvons exprimer que pour Gorthaur-EXP3 nous avons :

$$\forall f_b(t); \forall t \in [1; T]; \forall k > 0; \forall \eta [0; 1]; \forall b \in B$$

$$\begin{aligned} G_{max,f}(T) - \mathbb{E}(G_{Gorthaur,f}(T)) \\ \leq (e - 1)\eta G_{max,f}(T) + \frac{|\mathcal{B}| \ln(|\mathcal{B}|)}{\eta} \end{aligned}$$

où $G_{max,f}(T)$ est le gain maximum de fitness qui peut être obtenu à l'horizon T en sélectionnant systématiquement l'algorithme optimal qui correspond le mieux au compromis précision/diversité et $G_{Gorthaur,f}(T)$ est le gain de fitness obtenu avec l'algorithme Gorthaur-EXP3.

Analyse du compromis précision/diversité de Gorthaur-EXP3. En suivant scrupuleusement la preuve générale de l'algorithme EXP3 [Aue+02] (Section 3 de son article), nous pouvons démontrer la convergence du compromis précision/diversité de Gorthaur-EXP3 lorsqu'on utilise une sélection optimale d'algorithmes pour une valeur prédéfinie de $\Theta = \frac{\pi}{4}$ correspondant à un équilibre entre précision et diversité. Nous avons établi la Preuve 1 (sketch) du Théorème 1 ci-dessous.

Preuve 1 (Sketch) du Théorème 1.

Démonstration.

$\forall t \in [1; T]; \forall b \in \mathcal{B}$ avec $|\mathcal{B}| > 0$ (pour simplifier les notations, notons $z = |\mathcal{B}|$); $\forall f_b(t); \forall \eta \in]0; 1]$, soit $W_t = \sum_{i=1}^z w_{b_i}(t)$ la somme des poids de chaque algorithme. Notez que $W_1 = z$ puisque $\forall b \in \mathcal{B}$, nous avons pour $t = 1$, $w_b(1) = 1$ et $\forall b \in (B); p_b(t) = (1 - \eta) \frac{w_b(t)}{W_t} + \frac{\eta}{z}$ (Voir Algorithme 1).

Pour toute séquence d'algorithmes de sélection tirée par *Gorthaur-EXP3*, nous avons :

$$\begin{aligned} \frac{W_{t+1}}{W_t} &= \sum_{i=1}^z \frac{w_{b_i}(t) \exp\left(\frac{\eta}{z} \widehat{f}_{b_i}(t)\right)}{W_t}, \\ &= \sum_{i=1}^z \frac{p_{b_i}(t) - \frac{\eta}{z}}{1 - \eta} \exp\left(\frac{\eta}{z} \widehat{f}_{b_i}(t)\right), \end{aligned}$$

Comme $\widehat{f}_b(t) \leq \frac{1}{p_b(t)} \leq \frac{z}{\eta}$, et $\exp(x) \leq 1 + x + (e - 2)x^2$ quand $x \leq 1$, nous avons :

$$\frac{W_{t+1}}{W_t} \leq \sum_{i=1}^z \frac{p_{b_i}(t) - \frac{\eta}{z}}{1 - \eta} \left[1 + \frac{\eta}{z} \widehat{f}_{b_i}(t) + (e - 2) \left(\frac{\eta}{z}\right)^2 \widehat{f}_{b_i}(t)^2 \right],$$

Puisque $\sum_{i=1}^z p_{b_i}(t) \widehat{f}_{b_i}(t) = f_b(t)$ et $\sum_{i=1}^z p_{b_i}(t) \widehat{f}_{b_i}(t)^2 \leq \sum_{i=1}^z \widehat{f}_{b_i}(t)$, on peut écrire :

$$\frac{W_{t+1}}{W_t} \leq 1 + \frac{\left(\frac{\eta}{z}\right)}{1 - \eta} f_b(t) + \frac{(e - 2) \left(\frac{\eta}{z}\right)^2}{1 - \eta} \sum_{i=1}^z \widehat{f}_{b_i}(t),$$

En considérant que $1 + \Delta \leq \exp(\Delta)$, nous avons :

$$\ln\left(\frac{W_{t+1}}{W_t}\right) \leq \frac{\left(\frac{\eta}{z}\right)}{1 - \eta} f_b(t) + \frac{(e - 2) \left(\frac{\eta}{z}\right)^2}{1 - \eta} \sum_{i=1}^z \widehat{f}_{b_i}(t),$$

En sommant sur t (en opérant avec un horizon fini T), nous obtenons :

$$\ln\left(\frac{W_{T+1}}{W_1}\right) \leq \frac{\eta}{1 - \eta} G_{\text{Gorthaur},f}(T) + \frac{(e - 2) \left(\frac{\eta}{z}\right)^2}{1 - \eta} \sum_{i=1}^z \sum_{t=1}^T \widehat{f}_{b_i}(t),$$

Pour toute sélection d'algorithme b , nous avons :

$$\ln\left(\frac{W_{T+1}}{W_1}\right) \geq \frac{\ln(W_b(T+1))}{\ln(W_1)} = \frac{\eta}{z} \sum_{t=1}^T \widehat{f}_b(t) - \ln(z),$$

nous obtenons donc l'équation suivante :

$$G_{Gorthaur,f}(T) \geq (1 - \eta) \left(\sum_{t=1}^T \widehat{f}_b(t) - \frac{z \ln(z)}{\eta} \right) - (e - 2) \frac{\eta}{z} \sum_{i=1}^z \sum_{t=1}^T \widehat{f}_{b_i}(t),$$

En prenant l'espérance des deux côtés de l'équation précédente et comme b est choisi arbitrairement tel que $\sum_{t=1}^T f_{b_i}(t) \leq \eta G_{max}$, nous avons :

$$\mathbb{E} [G_{Gorthaur,f}(T)] \geq (1 - \eta) \left(G_{max,f}(T) - \frac{z \ln(z)}{\eta} \right) - (e - 2) \eta G_{max,f}(T),$$

et donc :

$$\mathbb{E} [G_{Gorthaur,f}(T)] - G_{max,f}(T) \geq -\eta (e - 1) G_{max,f}(T) - \frac{z \ln(z)}{\eta},$$

ce qui mène directement au Théorème 1 :

$$G_{max,f}(T) - \mathbb{E} [G_{Gorthaur,f}(T)] \leq (e - 1) \eta G_{max,f}(T) + \frac{|\mathcal{B}| \ln(|\mathcal{B}|)}{\eta}.$$

Pour plus d'informations sur l'analyse complète, à la fois notre preuve résumée et notre théorème s'appuient sur l'article clé [Aue+02] et sa preuve détaillée en Section 3.

Ainsi, comme *Gorthaur-EXP3* sélectionne les différents algorithmes de son portefeuille selon leur capacité à atteindre le compromis souhaité, on peut supposer qu'au pire, la borne supérieure du regret final de *Gorthaur-EXP3* sera la somme des bornes supérieures des regrets des algorithmes du portefeuille proportionnellement à leur taux de sélection à l'horizon T . Cependant, selon la Preuve 1 du Théorème 1, puisque *Gorthaur-EXP3* effectuera son étape d'exploitation en sélectionnant l'algorithme le plus optimal de son portefeuille, la borne supérieure du regret final de *Gorthaur-EXP3* tendra vers celle de son algorithme optimal.

À la sous-section suivante, nous étendons notre méthode au problème de tirage multiple d'éléments.

1.4.4 Méthodes *Gorthaur* à tirage multiple

1.4.4.1 Sélection d'algorithmes pour la recommandation multicritère de type top-k

Définition du problème. Afin d'optimiser plusieurs critères en ligne, nous avons proposé *Gorthaur-EXP3* [Gut+21] pour sélectionner les algorithmes de recommandation optimaux en rapport à ces critères et nous avons ensuite généralisé pour tout algorithme de *MAB* (*Gorthaur-MAB*) [Let+24].

Cependant, d'un point de vue plus large, l'un des enjeux de l'optimisation des systèmes de recommandation est de trouver un équilibre entre les préférences et les objectifs de plusieurs décideurs, y compris les utilisateurs et les acteurs impliqués dans le système (p.ex., le concep-

teur ou la plate-forme fournissant les éléments à recommander). Ainsi, lorsqu'il est possible de recommander plusieurs éléments à un utilisateur à un moment donné à l'aide de l'ensemble de recommandations S_t , le problème d'optimisation peut être formulé comme suit :

Soient \mathcal{U} l'ensemble des utilisateurs et \mathcal{V} l'ensemble des acteurs liés à un système de recommandation. L'objectif est de trouver les ensembles de k éléments recommandés qui satisfont les critères de tous les décideurs impliqués sur un horizon T . Étant donné qu'à chaque itération un seul utilisateur $u \in \mathcal{U}$ et $\mathcal{V}' \subseteq \mathcal{V}$ acteurs sont considérés, le problème de sélection d'un super-bras S à l'itération t peut être exprimé comme suit :

$$S_t = \operatorname{argmax}_{S \subseteq \mathcal{A}} F_u(S) \cap \bigcap_{v \in \mathcal{V}'} F_v(S)$$

où $F_u(S)$ et $F_v(S)$ représentent respectivement les fonctions d'évaluation des critères de l'utilisateur u et de l'acteur v , et S représente un ensemble de k éléments (bras) $a \in \mathcal{A}$ à recommander. Le système de recommandation vise à trouver l'ensemble optimal d'éléments S_t qui maximise la satisfaction de tous les utilisateurs et acteurs, tout en assurant que leurs objectifs possiblement conflictuels soient équilibrés.

Cependant, à mesure que le nombre de décideurs augmente, le problème d'optimisation devient plus complexe et la demande en termes de calcul augmente. Par conséquent, il peut être nécessaire d'utiliser des techniques d'optimisation plus adaptées, telles que des algorithmes d'optimisation multi-objectifs [Öne+18], pour résoudre efficacement le problème et trouver un ensemble optimal d'éléments à recommander qui satisferont aux mieux les critères de tous les décideurs.

À la lumière de cela, nous avons proposé les deux approches suivantes :

1. *Multiple-Play-Gorthaur*, qui vise à optimiser un compromis équilibré entre les critères des décideurs à chaque itération ;
2. *Budgeted-Gorthaur*, qui divise la recommandation S_t en plusieurs sous-ensembles d'éléments, chacun visant à optimiser différents compromis entre les critères.

1.4.4.2 Nouvelles approches : *Multiple-Play Gorthaur* et *Budgeted Gorthaur*

Similitudes et divergences avec les premières approches Gorthaur. *Multiple-Play-Gorthaur* (*MP-Gorthaur*) et *Budgeted-Gorthaur* possèdent une base commune aux méthodes Gorthaur à tirage simple en utilisant le même système de Compass [MS08] (voir Section 1.4.1) pour calculer le fitness des algorithmes. En revanche ces nouvelles variantes adoptent une approche à tirage multiple, ce qui signifie qu'elles peuvent sélectionner, selon la version utilisée (MP ou Budgeted) : soit un seul, soit plusieurs algorithmes permettant ensuite de recommander plusieurs éléments.

Motivations par rapport aux méthodes d'origines. Les motivations de ce travail diffèrent légèrement de celles de les propositions précédentes de *Gorthaur* [Gut+21], dans la mesure où

il vise à optimiser les critères (objectifs) des acteurs $v_i \in V$ tout en garantissant des minimums satisfaisants pour un utilisateur $u_t \in U$, ou inversement. Dans ces nouvelles méthodes, nous considérons un sous-ensemble d'éléments pour optimiser chaque critère indépendamment afin de maximiser les chances qu'une recommandation inclue au moins un élément satisfaisant pour chaque objectif.

Détails des contributions supplémentaires apportées par rapport aux méthodes d'origines. Les travaux que nous présentons dans cette partie sont les plus récents et ont mis un point d'orgue aux méthodes *Gorthaur*. Ils ont été publiés en juillet 2024 dans le journal *Expert Systems With Applications* [Let+24]. Ils englobent les contributions clés suivantes :

1. L'application de *Gorthaur-MAB* [Gut+21] à des scénarios multicritères impliquant des recommandations de k -éléments en utilisant l'algorithme de multiple-play [AVW87 ; LKC19]. Deux méthodes sont proposées : La première, appelée *Multiple-Play Gorthaur*, vise à optimiser un compromis entre les critères et peut être considérée comme une référence pour l'évaluation empirique proposée dans notre article [Let+24]. La seconde méthode, nommée *Budgeted-Gorthaur*, vise à maximiser les chances de soumettre des éléments satisfaisant chaque critère en considérant plusieurs compromis entre les critères lors de l'étape de sélection d'algorithmes.
2. Une étude approfondie de l'impact du méta-algorithme choisi sur les performances de l'approche. Plutôt que d'utiliser l'algorithme *EXP3* pour résoudre le problème de sélection d'algorithmes comme réalisé dans [Gut+21], nous avons expérimenté avec plusieurs algorithmes de *MAB* et *CMAB* en tant que méta-algorithme. Il en ressort que *Thompson Sampling* montre des résultats significativement meilleurs. C'est pourquoi nous proposons une preuve schématique afin de renforcer nos résultats empiriques par des garanties théoriques.

1.4.4.3 Méthode *Multiple-play Gorthaur*

Le fonctionnement de la méthode *MP-Gorthaur* est décrite à l'aide de l'algorithme 2. Cet algorithme présente un cadre général pour l'approche, qui doit donc être ajusté en fonction du méta-algorithme utilisé. Par conséquent, certaines parties du pseudo-code proposé peuvent nécessiter des ajustements en fonction de l'algorithme de *MAB* ou *CMAB* choisi comme méta-sélecteur. Le détail des ces ajustements est décrit dans notre article [Let+24].

Focus sur *MP-Gorthaur-TS*. Parmi les différentes approches *MAB* et *CMAB* étudiées comme sélecteur d'algorithmes (ϵ -*Greedy*, *LinTS*, *LinUCB*, *TS*, *UCB1*, *EXP3*), *MP-Gorthaur-TS* a montré les résultats les plus performants. En termes de précision globale, ses performances surpassent celles de l'utilisation exclusive de l'algorithme optimal pour l'horizon T lorsqu'on considère un portefeuille d'algorithmes *MAB* et *CMAB* dans le cas contextuel. Cela souligne

Algorithme 2 : Multiple-play Gorthaur avec partage des récompenses

Entrées : $x \in \mathcal{X}$: un ensemble de contextes utilisateurs. \mathcal{A} : L'ensemble des bras disponibles. \mathcal{B} : Portefeuille d'algorithmes. $\phi(S_t, Y_t)$: La stratégie pour considérer les retours des utilisateurs (voir Section 1.1.3.2). Θ : Angle exprimant le compromis désiré entre les objectifs considérés (voir Figure 1.4). k : nombre d'éléments par recommandation ($|S_t| = k$).

Initialisation : $\forall b_i \in \mathcal{B}$: Initialiser b_i .

```

1 for  $t = 1$  à  $T$  do
2   Considérer  $x_t \in X$  : un utilisateur et son contexte
3   for  $b_i \in \mathcal{B}$  do
4     Calculer les probabilités  $\mathbf{P}_{b_i,t}$  selon le méta-algorithme choisi
5   end
6   Sélectionner  $b_t$  selon la politique du méta-algorithme choisi et les probabilités
   ( $\mathbf{P}_{b_{1,t}}, \dots, \mathbf{P}_{b_{i,t}}$ ) ;
7    $S_t \leftarrow \emptyset$ 
8   while  $|S_t| < k$  do
9      $S_t = S_t \cup \{\operatorname{argmax}_{a \in \mathcal{A} \setminus S_t} \mathbb{E}[R_{t,a} \mid x_t]\}$ 
10  end
11  Observer les retours pour le super-bras  $S_t$  et calculer la récompense  $R_t = \phi(S_t, Y_t)$ 
12  for  $b_i \in \mathcal{B}$  do
13    Mettre à jour la politique  $\pi^{b_i}$  avec  $R_t$ 
14  end
15  Calculer  $acc(b_t, t) = \frac{\sum_{i=1}^t R(b_i, t)}{t_{b_t}}$ 
16  Calculer  $div(b_t, t) = 1 - \frac{C_v(N_{b_t, t})}{\sqrt{|\mathcal{A}|}}$ 
17  Calculer  $O_{b_t, t}^{norm} = (div^{norm}(b_t, t), acc^{norm}(b_t, t))$ 
18  Calculer  $f_{b_t, t} = |O_{b_t, t}^{norm}| \cos \alpha_{b_t, t} - \min_{b_i} \{|O_{b_i, t}^{norm}| \cos \alpha_{b_i, t}\}$ 
19  for  $b_i \in \mathcal{B}$  do
20    Mettre à jour la politique du méta-algorithme avec  $f_{b_i, t}$ 
21  end
22 end

```

la capacité de *MP-Gorthaur-TS* à s'adapter efficacement aux changements de méthode optimale à utiliser. Pour la suite nous nous concentrons donc sur la version de *Gorthaur* employant l'algorithme *Thompson Sampling*.

Étant donné que l'analyse théorique de *Thompson Sampling* repose fortement sur les seuils, la formulation des seuils présentée dans les travaux antérieurs de [AG13a] et [AG12] [AG13a] est adoptée ici et ajustée au cas spécifique de la sélection d'algorithmes. Nous complétons ainsi les définitions 10, 11, et 12 précédentes avec la définition suivante :

Définition 13. Seuils dans la sélection d’algorithmes Gorthaur-MAB. Selon la définition originale [AG13a], soient x_{b_i} et y_{b_i} deux seuils choisis pour chaque algorithme b_i tels que $\mu_{b_i} < x_{b_i} < y_{b_i} < \mu_{b^*}$. Soit $S_{c_{b_i}}(t) = \sum_{t'=1}^{t-1} r_{t'}$, avec $r_{t'} \in [0, 1]$, la somme des récompenses normalisées observées parmi les actions de l’algorithme b_i jusqu’à l’itération $t - 1$. Nous définissons $\hat{\psi}_{b_i}(t) = \frac{S_{c_{b_i}}(t)}{n_{b_i}(t)}$ où $\hat{\psi}_{b_i}(t) = 1$ lorsque $n_{b_i}(t) = 0$. Soit $E_{b_i}^\mu(t)$ l’événement où $\hat{\mu}_{b_i}(t) \leq x_{b_i}$ (c’est-à-dire l’événement où la moyenne estimée $\hat{\mu}_{b_i}$ n’est pas loin de μ_{b_i}), et soit $E_{b_i}^\psi(t)$ l’événement où $\hat{\psi}_{b_i}(t) \leq y_{b_i}$ (c’est-à-dire l’événement où le paramètre échantillonné n’est pas loin de μ_{b_i}). Enfin, nous définissons $L_{b_i}(T) = \frac{\ln T}{d(x_{b_i}, y_{b_i})}$.

Bornes du regret. Les bornes de regret pour les algorithmes de multiple-play ont déjà été établies pour tout algorithme de MAB dans des travaux antérieurs [AVW87]. Par conséquent, l’analyse est étendue pour inclure les bornes de regret spécifiquement pour *Gorthaur-TS*, ce qui permet d’identifier une instance appropriée de MAB/CMAB sur laquelle le *multiple-play* sera appliqué pour résoudre le problème de recommandation. Cette analyse se concentre donc sur le regret obtenu en utilisant des algorithmes sous-optimaux du portefeuille.

Théorème 2. Regret attendu de Gorthaur-TS du compromis précision/diversité dans un environnement stationnaire utilisant des priors Beta.

Soit \mathcal{B} l’ensemble des algorithmes du portefeuille de *Gorthaur-TS* et $b \in \mathcal{B}$ un algorithme spécifique du portefeuille où les algorithmes de \mathcal{B} sont associés à $|\mathcal{B}|$ bras dans ce problème de bandit stochastique à $|\mathcal{B}|$ bras.

Soit $f_b(t)$ la fitness de l’algorithme $b \in \mathcal{B}$ à l’itération $t \in [1, T]$ où T est l’horizon fini.

Alors, selon la preuve générale du regret attendu de TS [AG13a], nous pouvons exprimer qu’à l’horizon T pour *Gorthaur-TS*, nous avons :

$$\forall t \in [1, T]; \forall b \in \mathcal{B} \text{ avec } |\mathcal{B}| > 0; \forall f_b(t);$$

$$\mathbb{E}[\rho_{\text{Gorthaur},f}(T)] \leq (1 + \epsilon) \sum_{i=1}^{|\mathcal{B}|} \frac{\ln T}{d(\mu_{b_i}, \mu_{b^*})} \Delta_{b_i} + \mathcal{O}\left(\frac{|\mathcal{B}|}{\epsilon^2}\right)$$

où $d(\mu_{b_i}, \mu_{b^*}) = \mu_{b_i} \log \frac{\mu_{b_i}}{\mu_{b^*}} + (1 - \mu_{b_i}) \log \frac{(1 - \mu_{b_i})}{(1 - \mu_{b^*})}$,
 et $\Delta_{b_i} = \mu^* - \mu_{b_i}$ est le regret d’utiliser l’algorithme b_i au lieu de l’algorithme b^* ,
 et μ_{b_i} et Δ_{b_i} sont supposés constants pour $i \in [1, |\mathcal{B}|]$.

Pour résoudre le problème de sélection multicritère en utilisant l’algorithme *TS*, cette approche considère les récompenses $r_t \in [0, 1]$ obtenues par une fonction agrégative appliquée aux objectifs. Par conséquent, les bornes de regret fournies sont non triviales et nécessitent une analyse théorique pour assurer leur validité. Ainsi ci-dessous nous établissons la Preuve 2 (sketch) du Théorème 2.

Preuve 2 (Sketch) du Théorème 2.

Démonstration.

Selon la preuve générale de la convergence de *TS* [AG13a], le nombre attendu de fois où un algorithme sous-optimal $b_i \neq b^*$ est sélectionné peut être décomposé comme suit :

$$\begin{aligned} \mathbb{E}[n_{b_i}(T)] &= \sum_{t=1}^T \Pr(b_i(t) = b_i) \\ &= \sum_{t=1}^T \Pr(b_i(t) = b_i, E_{b_i}^\mu(t), E_{b_i}^\psi(t)) \end{aligned} \quad (1.8)$$

$$+ \sum_{t=1}^T \Pr(b_i(t) = b_i, E_{b_i}^\mu(t), \overline{E_{b_i}^\psi(t)}) \quad (1.9)$$

$$+ \sum_{t=1}^T \Pr(b_i(t) = b_i, \overline{E_{b_i}^\mu(t)}) \quad (1.10)$$

Dans cette preuve, chaque terme de l'équation ci-dessus doit être borné. Pour plus de commodité, ils seront chacun désignés par leur numéro d'Équations : (1.8), (1.9), et (1.10). En utilisant le Lemme 1 d'Agrawal [AG13a], l'Équation (1.8) peut être bornée comme suit (version courte) :

$$\begin{aligned} \sum_{t=1}^T \Pr(b_i(t) = b_i, E_{b_i}^\mu(t), E_{b_i}^\psi(t)) &= \sum_{t=1}^T \mathbb{E} \left[\mathbb{E} \left[I(b_i(t) = b_i, E_{b_i}^\mu(t), E_{b_i}^\psi(t)) \mid \mathcal{F}_{t-1} \right] \right] \\ &\leq \mathbb{E} \left[\left(\frac{1 - p_{b_i, t}}{p_{b_i, t}} \right) \Pr(b_i(t) = b^*, E_{b^*}^\mu(t), E_{b^*}^\psi(t) \mid \mathcal{F}_{t-1}) \right] \end{aligned}$$

où \mathcal{F}_{t-1} est un filtrage défini comme l'historique des sélections d'algorithmes dans \mathcal{B} jusqu'à l'itération $t - 1$ [AG13a] et $p_{b_i, t} = \Pr(\psi_1(t) > y_{b_i} \mid \mathcal{F}_{t-1})$. Notez que lorsque finalement l'algorithme b^* est exponentiellement choisi un nombre infini de fois, l'Équation (1.8) est bornée par la fonction de probabilité linéaire de jouer l'algorithme optimal (bras) (b^*). Cela signifie que :

$$\sum_{t=1}^T \Pr(b_i(t) = b_i, E_{b_i}^\mu(t), E_{b_i}^\psi(t)) \leq \mathcal{O}(1)$$

En appliquant le Lemme 3 et 4 [AG13a] et la Définition 13 pour borner l'Équation (1.9) nous obtenons (version courte) :

$$\sum_{t=1}^T \Pr(b_i(t) = b_i, E_{b_i}^\mu(t), \overline{E_{b_i}^\psi(t)}) = \sum_{t=1}^T \Pr(b_i(t) = b_i, n_{b_i}(t) \leq L_{b_i}(T), E_{b_i}^\mu(t), \overline{E_{b_i}^\psi(t)}) \quad (1.11)$$

$$+ \sum_{t=1}^T \Pr(b_i(t) = b_i, n_{b_i}(t) > L_{b_i}(T), E_{b_i}^\mu(t), \overline{E_{b_i}^\psi(t)}) \quad (1.12)$$

où l'Équation (1.12) est bornée par 1 et l'Équation (1.11) est bornée par $L_{b_i}(T) = \frac{\ln T}{d(x_{b_i}, y_{b_i})}$ étant donné que $\hat{\mu}_{b_i} \leq x_{b_i}$ (Voir Définition 13) ce qui résulte en un échantillonnage de $\psi > y_{b_i}$, avant que l'a posteriori ne se concentre autour de sa moyenne. Ainsi, nous obtenons :

$$\begin{aligned} \sum_{t=1}^T \Pr\left(b_i(t) = b_i, E_{b_i}^\mu(t), \overline{E_{b_i}^\psi(t)}\right) &\leq \frac{\ln T}{d(x_{b_i}, y_{b_i})} + 1 \\ &\leq \frac{\ln T}{d(\mu_{b_i}, \mu_{b^*})} (1 + \varepsilon)^2 + 1 \end{aligned}$$

Notez qu'il sera échantillonné $\frac{\ln T}{d(x_{b_i}, y_{b_i})}$ fois pour $\psi > y_{b_i}$ et avec une probabilité de $1/T$ une fois $\psi > y$. Cela signifie qu'après une exploration suffisante, cette partie de l'équation, représentant le coût de l'exploration, est attendue comme étant faible. Enfin :

$$\sum_{t=1}^T \Pr\left(b_i(t) = b_i, E_{b_i}^\mu(t), \overline{E_{b_i}^\psi(t)}\right) \leq \mathcal{O}\left(\frac{\ln T}{d(\mu_{b_i}, \mu_{b^*})}\right)$$

L'Équation (1.10) est bornée par Chernoff-Hoeffding. Ainsi, en utilisant la Définition 2 d'Agrawal [AG13a] et en appliquant leur Lemme 2 pour borner l'Équation (1.10), nous obtenons (version courte) :

$$\sum_{t=1}^T \Pr\left(b_i(t) = b_i, \overline{E_{b_i}^\mu(t)}\right) \leq \frac{1}{d(x_{b_i}, \mu_{b_i})} + 1$$

en utilisant l'inégalité de Pinsker, nous obtenons

$$\begin{aligned} \sum_{t=1}^T \Pr\left(b_i(t) = b_i, \overline{E_{b_i}^\mu(t)}\right) &\leq \frac{1}{d(x_{b_i}, \mu_{b_i})} + 1 \\ &\leq \mathcal{O}\left(\frac{1}{\varepsilon^2}\right) \quad (\text{puisque } x_{b_i} - \mu_{b_i} \geq \varepsilon \text{ (constant)}) \end{aligned}$$

Ainsi, pour un certain $\varepsilon \in [0, 1]$, soit $d(x_{b_i}, \mu_{b^*}) = \frac{d(\mu_{b_i}, \mu_{b^*})}{(1+\varepsilon)}$ et $d(x_{b_i}, y_{b_i}) = \frac{d(\mu_{b_i}, \mu_{b^*})}{(1+\varepsilon)^2}$ ce qui, en réunissant les 3 termes, conduit à :

$$\mathbb{E}[n_{b_i}(T)] \leq \mathcal{O}(1) + (1 + \varepsilon)^2 \frac{\ln T}{d(\mu_{b_i}, \mu_{b^*})} + \mathcal{O}\left(\frac{1}{\varepsilon^2}\right)$$

1.4.4.4 Méthode *Budgeted-Gorthaur*

Une autre approche intéressante pour résoudre les problèmes multi-objectifs est de considérer chaque objectif comme un problème de bandit-manchoit indépendant afin de trouver la solution optimale pour chaque objectif [Gab+11]. Cela transforme le problème multi-objectifs en un problème multi-bandits. Cette approche est particulièrement utile dans les contextes combinatoires où plusieurs bras sont recommandés à chaque tour et où différents objectifs peuvent être optimisés simultanément pour différents sous-ensembles de bras dans la recommandation.

Avec *Budgeted-Gorthaur*, nous proposons de nouveau une extension des travaux originaux [Gut+21]. Cette fois, les recommandations fournies par notre nouvelle approche sont construites en fusionnant un nombre $|k'|$ de sous-ensembles composés de k'_i éléments, tels que $\sum_i k'_i = k$, et où chaque sous-ensemble d'éléments est sélectionné pour répondre à un compromis spécifique entre les critères.

Plus précisément, pour chaque compromis i , l'approche identifie un algorithme optimal dans le portefeuille chargé de sélectionner k'_i éléments. Cela conduit à l'exploration de $|k'|$ parties différentes de l'ensemble de solutions Pareto. Pour ce faire, plusieurs méta-algorithmes, avec des angles de référence différents, sont employés. Étant donné que les angles de référence décrivent des compromis entre les critères, un même algorithme du portefeuille peut être très pertinent pour un méta-algorithme mais inadapté pour un autre, conduisant à un apprentissage différent pour chaque méta-algorithme. Les sous-ensembles de S_t déterminés par les algorithmes choisis par différents méta-algorithmes sont appelés "composantes" C_i .

L'algorithme 3 décrit le fonctionnement de *Budgeted-Gorthaur* dans un cadre général. Comme l'éventail des combinaisons potentielles de méta-algorithmes pouvant être utilisés est vaste, différents modèles peuvent nécessiter des ajustements uniques à l'algorithme 3 de la même manière que pour *MP-Gorthaur*. Le détail de ces ajustements est décrit dans notre article [Let+24].

Modification importante. Il est à noter une modification majeure par rapport à l'utilisation d'un seul méta-algorithme : les fonctions de mise à jour des méta-algorithmes impliquent que les valeurs de performance de l'algorithme $b_{i,t}$, $acc(b_{i,t}, t)$ et $div(b_{i,t}, t)$ soient mises à jour en fonction des performances observées sur la composante C_i et non sur la recommandation complète S_t . Ainsi, la performance de précision globale $acc(b_{i,t}, t)$ de l'algorithme $b_{i,t}$ sélectionné par le méta-algorithme M_i à l'itération t est déterminée selon l'équation suivante :

$$acc(b_{i,t}, t) = \frac{g(b_{i,t}, t)}{t_{b_{i,t}}} \quad (1.13)$$

où $t_{b_{i,t}}$ désigne le nombre d'itérations pendant lesquelles l'algorithme $b_{i,t}$ a été sélectionné par le méta-algorithme M_i , et :

$$g(b_{i,t}, t) = \sum_{v=1}^{k'_{i-1}+k'_i} \frac{R_v}{k'_i}$$

où $k'_0 = 0$ et $R_v \in \{0, 1\}$ est la récompense associée à l'élément v dans le super-bras S_t . De même, la métrique de diversité agrégée est mise à jour pour chaque composante. Ainsi, à chaque itération et pour chaque composante, la performance de diversité agrégée $div(b_{i,t}, t)$ est évaluée selon l'équation suivante :

$$div(b_{i,t}, t) = 1 - \frac{C_v(N_{b_{i,t}}, t)}{\sqrt{|\mathcal{A}|}} \quad (1.14)$$

Algorithme 3 : Budgeted-Gorthaur

Entrées : X : ensemble de contextes d'utilisateurs. \mathcal{A} : ensemble de bras (éléments) à recommander. \mathcal{B} : Portefeuille d'algorithmes. Θ : vecteur d'angles exprimant les compromis adaptés entre les objectifs considérés pour chaque composante. k : nombre d'éléments par recommandation (cardinalité de S_t). k' : vecteur des cardinalités des composantes de S_t .

Init : pour $i = \{1, \dots, |k'|\}$: $k'_i = \text{eq. A.6}$, initialiser le méta-algorithme M_i ; pour $j = \{1, \dots, |\mathcal{B}|\}$: initialiser l'algorithme b_j

```

1 for  $t = 1$  à  $T$  do
2   Considérer  $x_t \in X$  : un utilisateur et son contexte
3    $S_t \leftarrow \emptyset$ 
4   for  $i=1$  à  $|k'|$  : do
5     for  $b_j \in \mathcal{B}$  do
6       Calculer les probabilités  $\mathbf{P}_{i,b_j,t}$  selon le méta-algorithme  $M_i$ 
7     end
8     Sélectionner  $b_{i,t}$  selon la politique  $M_i$  et les probabilités  $(\mathbf{P}_{i,b_1,t}, \dots, \mathbf{P}_{i,b_j,t})$ 
9     while  $|S_t| < \sum_{v=1}^i k'_v$  do
10       $S_t = S_t \cup \{\text{argmax}_{a \in \mathcal{A} \setminus S_t}^{b_{i,t}} \mathbb{E}[R_{t,a} | x_t]\}$ 
11    end
12  end
13  Observer les retours pour le super-bras  $S_t$  et calculer  $R_t = \phi(S_t, Y_t)$ 
14  for  $b_j \in \mathcal{B}$  do
15    Mettre à jour la politique  $\pi^{b_j}$  avec  $R_t$  selon les besoins de  $b_j$ 
16  end
17  for  $i=1$  à  $|M|$  : do
18    Calculer  $\text{acc}(b_{i,t}, t) = \frac{g(b_{i,t})}{t_{b_{i,t}}}$ 
19    Calculer  $\text{div}(b_{i,t}, t) = 1 - \frac{C_v(N_{b_{i,t},t})}{\sqrt{|\mathcal{A}|}}$ 
20    Calculer  $O_{b_{i,t},t}^{\text{norm}} = (\text{div}^{\text{norm}}(b_{i,t}, t), \text{acc}^{\text{norm}}(b_{i,t}, t))$ 
21    Calculer  $f_{i,b_{i,t},t} = |O_{b_{i,t},t}^{\text{norm}}| \cos \alpha_{i,b_{i,t},t} - \min_{b_i} \{|O_{b_{i,t},t}^{\text{norm}}| \cos \alpha_{i,b_{i,t},t}\}$ 
22    for  $b_j \in \mathcal{B}$  do
23      Mettre à jour la politique du méta-algorithme  $M_i$  avec  $f_{i,b_j,t}$ 
24    end
25  end
26 end

```

Ici $N_{b_i,t} = \{n_{a_1}(b_i,t), \dots, n_{a_{|A|}}(b_i,t)\}$ où $n_{a_j}(b_i,t)$ représente le nombre de fois que l'élément a_j a été recommandé par l'algorithme b_i jusqu'à l'itération actuelle.

Ainsi, différents méta-algorithmes M_j et M_h peuvent observer des métriques de performance distinctes $acc(b_i,t)$ et $div(b_i,t)$, pour un même algorithme b_i dans le portefeuille.

Exemple de Budgeted-Gorthaur avec EXP3-754515. L'algorithme *Budgeted-Gorthaur* décrit un cadre généraliste. Dans l'Annexe A.2 nous donnons un exemple d'implémentation possible de *Budgeted-Gorthaur : EXP3-754515*, comme nous l'avons proposé dans [Let+24] (c.-à-d., trois instances d'EXP3 sont utilisées comme méta-algorithmes).

1.4.5 Synthèse des résultats et discussion sur les méthodes Gorthaur

En résumé. Dans le cadre du tirage simple *Gorthaur-EXP3* surpasse *Gorthaur-Wheel* mais peut manquer de réactivité en environnement non-stationnaire (covariate shift). Dans le cadre du tirage multiple, les résultats obtenus démontrent que les méthodes *MP-Gorthaur-TS* et *MP-Gorthaur-EXP3* émergent comme les plus performantes parmi les variantes *Gorthaur*, en raison de leur capacité à bien équilibrer la précision globale et la diversité, ainsi qu'à s'adapter aux changements de stationnarité. *Budgeted-Gorthaur* est compétitif mais mieux adapté pour des applications nécessitant l'optimisation de plusieurs compromis entre des critères.

Dans cette section, nous proposons une synthèse et une discussion pour mieux appréhender ces résultats ainsi que les différences, avantages et inconvénients de la suite de méthodes *Gorthaur* que nous avons développée. Néanmoins, le détail des résultats est disponible dans les articles [Gut+19b ; Gut+21 ; Let+24].

La suite de méthodes *Gorthaur* inclut plusieurs variantes conçues pour s'adapter à différents contextes et exigences de recommandation. Parmi ces méthodes, on retrouve :

- **Gorthaur-Wheel** [Gut+19b] - Cette méthode utilise une stratégie de sélection par roulette proportionnelle (*roulette wheel*) pour choisir l'algorithme le plus optimal dans un portefeuille. Elle vise à maximiser un compromis entre précision et diversité selon des pondérations prédéfinies. Cependant, cette approche peut souffrir d'une perte de précision globale dans certaines situations notamment en cas de changement de régime.
- **Gorthaur-EXP3** [Gut+21] - Pour améliorer les performances et surmonter les limitations de *Gorthaur-Wheel*, nous avons développé *Gorthaur-EXP3* en utilisant l'algorithme *EXP3* (*Exponential-weight algorithm for Exploration and Exploitation*). Ce modèle est particulièrement efficace pour résoudre les problèmes de recommandation dans des environnements non stationnaires et dynamiques, car il équilibre l'exploration continue et l'exploitation de l'algorithme optimal. Il a aussi l'avantage de donner des garanties théoriques solides.
- **Multiple-play Gorthaur** [Let+24] - *MP-Gorthaur* généralise la méthode *Gorthaur* pour les recommandations de plusieurs éléments à chaque itération (top-k recommandations). Cette extension de *Gorthaur* n'est pas limitée à un méta algorithme exclusif. Cependant, nous montrons empiriquement que l'usage de *Thompson Sampling* (*TS*)

offre une précision globale significativement supérieure aux approches concurrentes. Cette amélioration est notamment due à une capture plus efficace du changement d'optimalité parmi les algorithmes de recommandation à l'aide d'un compromis plus efficace entre exploration et exploitation des algorithmes de recommandation. Pour cette approche, nos travaux fournissent également des garanties théoriques. En revanche, *MP-Gorthaur-TS* a une complexité computationnelle plus élevée.

- **Budgeted Gorthaur** [Let+24] - Plutôt que de traiter le problème comme un seul processus de bandit-manchot (*Multi-Armed Bandit - MAB*), cette approche considère chaque objectif comme un problème de *MAB* distinct. Cela permet de trouver des solutions pour plusieurs compromis entre les objectifs au travers de sous-ensembles d'éléments optimisés par compromis. *Budgeted Gorthaur* propose ainsi une approche plus flexible pour atteindre un compromis adéquat entre les critères et élargir la couverture de l'ensemble de Pareto des solutions optimales. Il est conçu pour surmonter la limitation naturelle des méthodes a priori qui ne parviennent pas à une couverture uniforme de l'ensemble Pareto. Son avantage est de permettre d'explorer plusieurs compromis entre les critères et offre plus de flexibilité dans les solutions de recommandation. Comparé à *Multiple-play Gorthaur*, il permet de répondre aux besoins de divers cas d'utilisation.

En conclusion, ces méthodes visent toutes à optimiser simultanément plusieurs critères (dans nos travaux : précision et diversité) dans les systèmes de recommandation en ligne, et chacune d'entre elles offre des avantages particuliers en fonction du contexte d'application et des exigences des utilisateurs. En synthèse un tableau récapitulatif est fourni pour mieux les appréhender et les comparer (Voir Tableau 1.1).

Méthode	Stratégie de sélection	Recommandations	Avantages	Inconvénients
<i>Gorthaur-Wheel</i>	Roulette Proportionnelle	Simple : Sélectionne un seul algorithme par itération pour recommander un seul élément.	Facile à mettre en œuvre ; Offre une approche équilibrée entre précision et diversité.	Peut manquer de précision dans des contextes spécifiques ; Convergence lente vers l'algorithme optimal.
<i>Gorthaur-EXP3</i>	Bandit-manchot avec <i>EXP3</i>	Simple : Sélectionne un seul algorithme par itération pour recommander un seul élément.	Bonne adaptation à des environnements non-stationnaires ; Exploration dynamique des algorithmes.	Peut échouer à capturer la diversité optimale ; Non conçu pour les recommandations multiples.
<i>MP-Gorthaur</i>	<i>Multiple-play</i> avec <i>MAB/CMAB</i>	Multiple : Sélectionne un seul algorithme recommandant plusieurs éléments.	Optimise encore mieux précision et diversité ; Meilleure adaptation à la non-stationnarité.	Complexité computationnelle élevée ; Nécessite une gestion fine des interactions entre recommandations.
<i>Budgeted-Gorthaur</i>	Multi-bandit : Plusieurs <i>EXP3</i> chaque composant pour un compromis.	Multiple : Sélectionne plusieurs algorithmes optimaux qui recommandent chacun plusieurs éléments.	Grande flexibilité pour couvrir l'ensemble de Pareto des solutions ; Adapte les compromis selon les besoins.	Complexité plus élevée pour la gestion des budgets et des angles de référence ; Réglage des paramètres nécessaire.

TABLE 1.1 – Comparaison des différentes variantes de la méthode *Gorthaur*

1.4.6 Réflexion critique

Les méthodes d'« Intelligence Artificielle », qu'elles reposent sur des modèles ou sur l'intelligence computationnelle, partent du principe que tout algorithme prouvé théoriquement pourrait être infaillible dans ses prédictions, à condition de respecter les limites établies par cette preuve. Cependant, lorsqu'il s'agit d'aide à la décision humaine ou de systèmes de recommandation, le facteur humain doit impérativement être pris en compte. En effet, l'être humain est une somme d'équations parfaites qui, malgré tout, engendrent un être fondamentalement imparfait.

Ainsi, il devient essentiel de développer des approches dont l'objectif ne se limite pas à atteindre une valeur « juste » de précision, mais qui cherchent également à anticiper et à compenser les nombreux facteurs humains auxquels les algorithmes sont soumis lors de leur entraînement. Ces facteurs incluent les biais, la parcimonie des données, ainsi que les changements de préférences ou d'opinions. C'est en partie dans cette optique que les méthodes décrites précédemment ont été développées : la prise en compte de l'humain au centre des décisions purement mathématiques.

OPTIMISATION MULTICRITÈRE POUR LA SÉLECTION D'ATTRIBUTS MÉDICAUX

Simplicity versus Complexity.

« *Simplicity does not precede complexity, but follows it.* » [Per82]

Alan Perlis - 1982

Sommaire

2.1 Problème de sélection d'attributs multi-objectif	84
2.2 Positionnement de nos contributions	91
2.3 Genetic Algorithm with Multi-Objective Compass - <i>GAwC</i>	92

Ce chapitre aborde le problème de la sélection d'attributs, une problématique centrale en apprentissage automatique et en intelligence artificielle, particulièrement pertinente pour les tâches de classification en haute dimension (c'est-à-dire lorsque le nombre d'attributs est élevé). L'objectif principal est d'identifier un sous-ensemble $\mathcal{S} \subseteq \mathcal{V}$ d'attributs qui permette de maintenir, voire d'améliorer, la performance du modèle d'apprentissage. Ce problème, reconnu comme NP-difficile [AK98], revêt une nature multi-objectifs [Jia+23 ; Kar+15 ; Li+20], où chaque critère d'optimisation peut être considéré comme un problème distinct avec ses propres solutions optimales.

La performance d'un modèle d'apprentissage peut dépendre de critères variés, dont l'importance peut varier selon l'application envisagée. Parmi ces critères peuvent figurer par exemple : l'exactitude (*accuracy*), la précision, le *F-Score*, l'*AUC* (témoignant de la capacité prédictive), ou encore le coefficient de corrélation de Matthews (*MCC*) [CTJ21]. Dans le cas spécifique de ce chapitre, l'objectif consistait dans le cadre de la classification binaire dans le domaine médical, à maintenir une *AUC* aussi élevée que possible tout en limitant au maximum le nombre d'attributs médicaux pris en compte. Bien que l'exactitude (*accuracy*) ait été un critère secondaire, elle restait importante pour évaluer et comparer les performances avec des approches concurrentes issues de la littérature. Cette demande de l'utilisateur qui était dans notre cas un PU-PH (Professeur des Universités - Praticien Hospitalier), peut se traduire par un front Pareto qui correspond à l'ensemble des compromis optimaux entre les différents critères. La performance vue par l'utilisateur est donc ici la capacité qu'ont des solutions proposées à mieux correspondre au compromis souhaité entre l'*AUC*, le nombre d'attributs et l'exactitude. Ainsi, une fois un sous-ensemble d'attributs sélectionné, sa pertinence est ensuite évaluée en l'uti-

lisant comme ensemble de variables prédictives dans une tâche d'apprentissage supervisé donnée (p.ex. classification) et en mesurant la performance obtenue.

Enfin, au-delà des considérations purement axées sur la performance, l'identification des attributs clés présente un avantage supplémentaire : elle rend le modèle plus facilement interprétable. Cet aspect est crucial, voire indispensable, dans des applications sensibles comme celles du domaine médical.

En raison de la nature potentiellement conflictuelle des objectifs — minimiser le nombre d'attributs tout en maximisant la précision et d'autres métriques de performance — la sélection d'attributs multi-objectif peut devenir un problème complexe qui nécessite des techniques d'optimisation avancées. Ainsi, dans ce chapitre, nous rappelons différentes approches de ce domaine, notamment les algorithmes évolutionnaires multi-objectifs (*MOEAs*), qui ont démontré leur efficacité pour équilibrer plusieurs objectifs en même temps. Nous présentons également diverses méthodes de sélection d'attributs multi-objectif, avec une évaluation de leurs avantages et inconvénients. Ensuite, nous introduisons notre contribution : **Genetic Algorithm with Multi-objective Compass** (*GAWC*) [Gut+22], une méthode novatrice combinant un algorithme génétique avec une approche de calcul de fitness basée sur un système de Compass [MS08] multicritère. L'algorithme *GAWC* est conçu pour optimiser simultanément plusieurs critères — c.-à-d., exactitude (*accuracy*), *AUC*, nombre d'attributs — dans le contexte de la classification médicale, où la sélection d'attributs identifiables et explicables est requise. Il s'inspire directement de l'algorithme *MOEA/D* [ZL07].

Dans ce chapitre, nous proposons un aperçu parmi les nombreuses méthodes existantes et nous présentons les contributions de notre nouvelle approche, en soulignant son efficacité et sa pertinence pour des applications concrètes, en particulier dans le domaine médical.

Les travaux qui en sont issus font référence à ma collaboration avec le CHU d'Angers et l'ESEO de 2021 à 2022 (Pr Pierre Abraham (PU-PH) - CHU d'Angers, Daniel Schang et Olivier Camp - ESEO). De ce fait, ces travaux s'inscrivent avant tout dans une démarche appliquée, répondant à des besoins spécifiques en médecine, et s'appuient sur des concepts issus de l'optimisation multi-objectif pour proposer des solutions adaptées. Leur portée est principalement pratique et centrée sur les enjeux applicatifs du domaine médical.

Note. En anglais, il n'y a généralement qu'un seul terme, « *features* » ou « *dimension* » pour évoquer ce qu'en français nous pouvons dénommer p.ex., variables de contexte, caractéristiques, attributs, variables d'entrée, dimensions, informations contextuelles. De ce fait, pour plus de clarté et d'uniformité du vocabulaire, je fais référence au terme d'*attributs* en français pour évoquer le terme de « *features* » en anglais.

Contexte des contributions de ce chapitre

Historique. Ce chapitre et les travaux qui en résultent sont liés à mes recherches en collaboration avec le CHU d'Angers et l'ESEO. Il répond à la fois à mon intérêt pour le domaine médical et à ma volonté de monter en compétences sur l'axe Métaheuristique et Optimisation Combinatoire (MOC) du LERIA.

Équipe. J'ai travaillé avec le Pr Pierre Abraham (CHU d'Angers, INSERM 1083, CNRS 6015) et avec mon ancien co-encadrant de thèse Olivier Camp (ESEO) et mon ancien collègue également enseignant-chercheur Daniel Schang (ESEO).

Environnement et organisation. Une partie de ces travaux se sont déroulés durant la seconde partie de la pandémie de *COVID-19* impliquant un travail de recherche à la fois en distanciel et en présentiel. Des réunions de collaboration se déroulaient au fil de l'eau mais nous avions un point récurrent systématique au moins une fois par semaine.

Mon implication et mes contributions. J'ai assuré les missions suivantes pour ce travail de recherche : formalisation, conception, implémentation de la partie calcul du fitness et refonte de l'ensemble du code, expérimentation, écriture/revue de l'article et choix du journal dans lequel publier.

Note importante. Les sections de ce chapitre s'appuient sur les travaux suivants : [Gut+22].

Ces travaux ont été réalisés en collaboration avec les co-auteurs, et les contributions intellectuelles sont partagées entre eux. Ainsi, certaines sections de ce chapitre peuvent provenir de traductions, de reformulations ou encore de formulations textuelles de ces travaux originaux.

2.1 Problème de sélection d'attributs multi-objectif

Cette section propose un aperçu synthétique de l'état de l'art concernant le problème de la sélection d'attributs multi-objectif. Elle s'appuie en grande partie sur les éléments développés dans notre article [Gut+22].

Le problème de sélection d'attributs (features) multi-objectif fait écho aux besoins de diminuer le nombre de dimensions sur lequel repose un apprentissage automatique. Plus connu sous le nom de fléau de la dimensionnalité (Curse of Dimensionality), la complexité et la performance d'un apprentissage repose non seulement sur le volume de données disponibles mais également sur le nombre de dimensions – attributs – à prendre en compte. Les données et les attributs déterminent ainsi la limite supérieure des performances de l'apprentissage, tandis que les modèles et les algorithmes tentent simplement de s'en approcher. Dans ce contexte, il devient parfois incontournable de diminuer le nombre de dimensions soit par des méthodes de réduction de dimensions, soit par des méthodes de sélection des attributs les plus pertinents.

De plus, des attributs de qualité fournissent des informations sémantiques et structurelles des données, ce qui permet à des modèles ordinaires d'obtenir de meilleures performances

d'apprentissage [KS17 ; Li+20 ; Wu+15 ; Yan+18]. C'est pourquoi selon la tâche d'apprentissage et l'« explicabilité » qu'on en attend, il devient plus pertinent de choisir la sélection d'attributs plutôt que la réduction.

La sélection d'attributs joue un rôle crucial dans la procédure de prétraitement des données pour l'apprentissage automatique. Contrairement à la réduction de dimension, qui transforme les attributs originaux en de nouveaux, la sélection permet de conserver les informations d'origine en modifiant uniquement la combinaison des attributs pour trouver des sous-ensembles ayant un fort pouvoir discriminant [Li+20 ; MX17 ; Yon+19].

Nos travaux se sont focalisés sur le domaine médical où il est crucial de pouvoir reposer sur des attributs identifiables, connus et interprétables tout au long du processus d'apprentissage. C'est pourquoi nous avons rapidement opté pour des méthodes de sélection plutôt que de réduction. De plus, le professeur avec qui j'ai collaboré en médecine, le Pr Pierre Abraham, avait pour forte contrainte de devoir reposer sur un nombre d'attributs certes moindre mais conservant à la fois tout leur pouvoir discriminant et explicable au sens médical. Outre ces contraintes humaines, il est à noter que l'approche de sélection d'attributs, en plus de bien entendu réduire le nombre de dimensions, présente des avantages tels que la prévention du sur-apprentissage, l'amélioration de la généralisabilité des modèles et le renforcement de la compréhension des relations entre attributs. Ce dernier point étant lui aussi également très important dans le domaine médical (p.ex., déterminer les covariables et interactions synergiques).

La sélection d'attributs multicritère sur laquelle nos travaux se sont focalisés, cherche à équilibrer des objectifs souvent conflictuels, tels que très souvent maximiser la précision tout en minimisant le nombre d'attributs, ce qui pose des problématiques en termes de recherche de compromis.

Cette section se concentrera dans un premier temps sur des approches plus générales de ce domaine, telles que les algorithmes évolutionnaires multi-objectifs (MOEAs) [Deb+02 ; Ma+23] qui ont démontré leur efficacité dans la recherche de solutions équilibrées entre plusieurs objectifs. Dans un second temps, je présenterai plus en détail certains algorithmes représentatifs développés spécifiquement pour la sélection d'attributs multi-objectif [Kar+15 ; Li+20]. Leur choix est justifié par leur pertinence et leur adoption dans des travaux similaires, et ils serviront de méthodes compétitrices pour évaluer la nouvelle approche que j'ai proposée.

2.1.1 Rappel de l'état de l'art sur l'optimisation multi-objectif

Dans le domaine de l'optimisation multi-objectif, de nombreuses approches ont été développées, notamment les algorithmes évolutionnaires et les méthodes de recherche heuristique. Parmi les algorithmes évolutionnaires multi-objectifs (MOEAs), certains des plus utilisés incluent le *NSGA-II* [Deb+02 ; EF09 ; Ma+23] (*Non-dominated Sorting Genetic Algorithm II*), le *NSGA-III* [YXW14], le *PESA-II* (*Pareto Envelope based Selection Algorithm*) [GSP16], le *SPEA2* (*Strength Pareto Evolutionary Algorithm 2*) [Zit01], ainsi que le *MOEA/D* (*Multi-Objective Evolutionary Algorithm based on Decomposition*) [ZL07]. Une autre classe notable est celle des algorithmes basés sur l'hypervolume, tels que le *IBEA* [ZK04], ainsi que son ex-

tension *IBMOLS* [BB07], le *SMS-EMOA* [BNE07], et le *HypE* [BZ11], qui utilisent des stratégies maximisant directement l'espace couvert par le front de Pareto pour garantir la convergence et la diversité au sein des solutions obtenues.

Des recherches récentes [Cos+24; Lie+19] ont mis en lumière l'importance des caractéristiques des paysages multi-objectifs et des objectifs hétérogènes dans la performance des algorithmes. Ces caractéristiques incluent la rugosité (k -interactions), la corrélation entre objectifs (ρ), ou encore la diversité et la distribution des optima locaux. Une compréhension approfondie de ces propriétés permet non seulement d'expliquer la performance observée des algorithmes sur différents problèmes, mais aussi de prédire leur comportement et de guider le choix algorithmique.

En particulier, l'hétérogénéité entre objectifs, analysée dans [Cos+24], joue un rôle crucial dans la difficulté d'un problème bi-objectifs. Les résultats montrent que les algorithmes comme le *MOEA/D* bénéficient particulièrement de la décomposition dans les paysages présentant des corrélations fortes (ρ), tandis que les approches basées sur l'hypervolume, comme le *SMS-EMOA*, sont mieux adaptées aux paysages multi-modaux avec un grand nombre d'optima locaux. Ces observations sont essentielles pour concevoir des algorithmes robustes et efficaces.

Afin d'illustrer certaines approches populaires en optimisation multi-objectifs, nous avons sélectionné certains algorithmes représentatifs pour en présenter les caractéristiques selon des critères tels que la méthode de sélection, la gestion de la diversité, ainsi que leurs avantages et inconvénients (voir Tableau 2.1).

Après avoir présenté différents algorithmes évolutionnaires multi-objectifs (MOEAs) dans le tableau 2.1, il est important de souligner qu'il existe des plateformes conçues pour évaluer la performance de ces algorithmes, comme *ParadisEO-MOEO* [Lie+07]. Cette plateforme permet non seulement de développer des algorithmes multi-objectifs, mais aussi de gérer les archives de solutions non dominées, d'évaluer leurs performances à l'aide de divers indicateurs, et d'intégrer des mécanismes de parallélisme et d'hybridation pour optimiser l'exécution sur des systèmes à grande échelle. *ParadisEO-MOEO* est donc un outil permettant de tester et analyser l'efficacité d'algorithmes comme ceux présentés dans le tableau 2.1. Il permet ainsi aux chercheurs de comparer leurs performances sur divers problèmes et d'identifier les approches les plus adaptées aux caractéristiques spécifiques de chaque paysage multi-objectif.

Vers la sélection d'attributs multi-objectif. Nos travaux de recherche ont porté sur une application plus spécifique de l'optimisation multi-objectifs : la sélection d'attributs multi-objectifs (*MOFS Multi-Objective Features Selection*) pour la classification. Avant de proposer une contribution novatrice, nous nous sommes tout d'abord penchés sur les approches *MOFS*. Elles utilisent des algorithmes évolutionnaires multi-objectifs (*MOEAs*) développés spécifiquement pour identifier les sous-ensembles d'attributs les plus pertinents tout en équilibrant des objectifs conflictuels, tels que p.ex., maximiser la précision et minimiser le nombre d'attributs.

Cependant, la sélection d'attributs présente deux problématiques majeures. La première est le nombre important de combinaisons possibles de sous-ensembles, rendant impossible

Algorithme	Méthode de Sélection	Gestion de la Diversité	Avantages	Inconvénients
PESA-II	Division et sélection de sous-régions dans l'espace de recherche	Contrôle de densité dans une sous-région d'hypercubes	Bonne diversité dans des régions peu peuplées	Coût élevé dans le cas où de nombreux objectifs ont été définis
NSGA-II	Tri par dominance et distance d'encombrement	Distance d'encombrement pour les solutions du même rang	Simple, efficace, robuste	Difficulté dans des espaces à haute dimension et corrélations faibles
NSGA-III	Tri par dominance et points de référence	Utilisation de points de référence pour la distribution des solutions	Efficace pour des problèmes avec plusieurs objectifs	Dépend de la sélection des points de référence, difficile avec corrélations fortes
SPEA2	Sélection basée sur la dominance	Distance du k-plus-proche-voisin	Bonne convergence et diversité	Coût de calcul élevé avec une grande archive
MOEA/D	Décomposition en sous-problèmes scalaires	Voisinage basé sur les pondérations	Adapté aux problèmes fortement corrélés et nombreux objectifs	Sensible à l'hétérogénéité et au choix de la décomposition
SMS-EMOA	Élimination basée sur la contribution hypervolume	Maximisation de l'hypervolume	Bonne diversité et convergence	Coût élevé pour des espaces à haute dimension
IBEA	Indicateur basé sur l'hypervolume ou l'épsilon-dominance	Maximisation de l'indicateur	Flexible, applicable avec divers indicateurs	Sensible au choix de l'indicateur, coûteux pour des paysages complexes
HypE	Estimation stochastique de l'hypervolume	Échantillonnage probabiliste	Adapté aux hautes dimensions	Précision dépendante du nombre d'échantillons

TABLE 2.1 – Comparaison d'algorithmes évolutionnaires multi-objectifs (MOEAs).

le calcul de toutes les solutions possibles (problème NP-difficile [AK98]). La seconde est que différentes combinaisons d'attributs peuvent produire des résultats similaires, rendant difficile la distinction des combinaisons réellement efficaces [Kom+17 ; Zha+16].

Enfin, bien que nos travaux se soient concentrés sur une comparaison ciblée des méthodes de sélection d'attributs pour un même modèle d'apprentissage, il est crucial de noter que la difficulté de la sélection d'attributs est étroitement liée au modèle de *Machine Learning* employé, comme l'ont mis en évidence des travaux récents [LTV24]. Ces recherches, basées sur l'analyse des réseaux d'optima locaux (LON), montrent que la structure des paysages d'optimisation, influencée par le modèle utilisé, impacte significativement la capacité des algorithmes à converger efficacement vers des solutions optimales. Cette perspective souligne la nécessité de développer des approches adaptées aux spécificités des données, un point central pour la recherche en sélection d'attributs multi-objectifs.

Pour faire face à ces problématiques, différentes méthodes de recherche ont été proposées :

1. **Recherche complète** - Méthode qui explore toutes les solutions possibles mais qui est computationnellement coûteuse et inadaptée aux grands ensembles d'attributs. Exemples d'algorithmes : algorithme de recherche (parcours) en profondeur, algorithme A, A*. Ces algorithmes garantissent de trouver la solution optimale, mais leur complexité les rend inadaptés aux problèmes de grande taille.

2. **Recherche stochastique** - Méthode qui repose sur des processus aléatoires pour explorer l'espace de solutions, en ajoutant ou supprimant des attributs de manière probabiliste. Ces méthodes peuvent tomber dans des optima locaux et manquer certaines solutions potentielles, bien qu'elles soient moins coûteuses que la recherche complète. Exemples d'algorithmes : méthodes de Monte Carlo.
3. **Recherche métaheuristique** - Méthode qui repose sur des stratégies intelligentes d'exploration de l'espace des solutions. Elle inclut souvent l'utilisation de populations d'individus et des mécanismes comme le croisement et la mutation, mais aussi d'autres techniques comme la recherche locale, le recuit simulé ou les colonies de fourmis. Ces méthodes sont conçues pour éviter les optima locaux en explorant plus largement l'espace de recherche, souvent par une exploration non exhaustive mais efficace. Les algorithmes évolutionnaires font partie de cette catégorie, car ils utilisent des mécanismes génétiques basés sur des populations pour explorer l'espace de solutions et trouver des solutions optimales. D'autres approches, comme le recuit simulé et les colonies de fourmis, font également partie des métaheuristiques, mais elles emploient des stratégies différentes pour l'exploration de l'espace de recherche.

Nos travaux se sont plutôt focalisés sur la troisième famille de méthodes : « Recherche métaheuristique ». Ainsi, à la sous-section suivante, nous nous concentrons à rappeler les méthodes de métaheuristique les plus populaires pour la sélection d'attributs multi-objectif.

2.1.2 Comparaison de méthodes de type métaheuristique populaires en sélection d'attributs multi-objectif.

2.1.2.1 Méthodes de sélection

La recherche métaheuristique dans le cadre de la sélection d'attributs est divisée en trois principales méthodes de sélection : les méthodes de filtrage (*filter*), les méthodes d'enveloppe (*wrapper*), et les méthodes intégrées (*embedded*). Les méthodes *filter* sont indépendantes du modèle. Elles utilisent des objectifs indirects pour sélectionner des sous-ensembles. Les méthodes *wrapper* quant à elles traitent la sélection d'attributs comme un processus qui nécessite l'entraînement du modèle de prédiction pour chaque sous-ensemble. Enfin, les méthodes *embedded* intègrent la sélection d'attributs directement dans le processus d'entraînement du modèle.

2.1.2.2 Principales fonctions d'évaluation

Dans le cadre de la sélection d'attributs multicritères, plusieurs fonctions d'évaluation sont couramment utilisées pour guider le processus d'optimisation [Jia+23]. Parmi celles-ci, on peut rappeler :

- *Feature subset size* : Reflète la cardinalité du sous-ensemble sélectionné, utilisée comme un critère de contrôle de la parcimonie.

- *Classification accuracy* : Évalue la performance des modèles d'apprentissage supervisé basés sur les attributs sélectionnés.
- *Relevance* : Mesure la pertinence des attributs par rapport aux classes cibles, souvent à l'aide de métriques comme p.ex. les coefficients de corrélation.
- *Redundancy* : Quantifie les chevauchements d'information entre les attributs, permettant de limiter les duplications inutiles.

2.1.2.3 Choix des méthodes comparées

Dans cette sous-section, nous proposons de présenter et d'évaluer certaines méthodes issues de la littérature, pertinentes pour le cadre de la sélection multi-objectif d'attributs pour la classification notamment dans le domaine médical.

Nous avons choisi ces méthodes en raison de leur capacité à optimiser simultanément l'exactitude (*accuracy*) du modèle, à réduire la redondance des attributs et à explorer efficacement l'espace de recherche. Ainsi, les méthodes comparées ont été choisies pour situer notre méthode *GAWC* par rapport à des approches assez similaires :

- Ces méthodes ciblent des critères similaires, notamment la réduction des attributs (*Feature subset size*) et l'optimisation de l'exactitude (*classification accuracy*).
- Elles ont été validées dans des domaines analogues (classification médicale).
- Elles couvrent les familles *filter*, *wrapper*, et hybrides pour offrir un panorama plus représentatif d'approches métaheuristiques bien que notre méthode soit plutôt de type *wrapper*.
- L'inclusion de méthodes *filter* permet de mettre en évidence les compromis entre coût computationnel et précision.

Dans le contexte de la *MOFS*, plusieurs méthodes ont démontré leur efficacité dans divers domaines d'application [Jia+23]. Parmi elles, les approches basées sur des algorithmes évolutionnaires multi-objectifs sont particulièrement performantes, p.ex., la méthode *QEISS* (*Quasi-Equally Informative Subset Selection*) [Kar+15] ou l'algorithme *DMEA-FS* (*Dividing-based Many-objective Evolutionary Algorithm for Feature Selection*) [Li+20]. Ces méthodes, principalement de type *wrapper*, ont été appliquées avec succès dans des domaines tels que la classification médicale, la reconnaissance de formes, la détection d'anomalies ou encore la reconnaissance de caractères. Dans le Tableau 2.2, nous examinons plus spécifiquement les méthodes avec lesquelles nous nous sommes empiriquement comparées [Li+20] par la suite, en fonction de leur méthodes de sélection et de leur description, de leurs objectifs principaux, ainsi que de leurs avantages et inconvénients.

Méthode	Références	Description et méthodes de sélection	Objectifs principaux	Avantages	Inconvénients
W-DMEA-FS	[Li+20]	<i>Dividing-based Many-Objective Evolutionary Algorithm (DMEA-FS)</i> utilisant une méthode de sélection <i>wrapper</i> .	Maximiser la précision et minimiser le nombre d'attributs.	Haute précision et flexibilité d'intégration d'objectifs multiples.	Coût de calcul élevé dû aux évaluations répétées des modèles.
F-DMEA-FS	[Li+20]	Variante de <i>DMEA-FS</i> utilisant une méthode de sélection <i>filter</i> pour réduire le coût de calcul en utilisant des critères indépendants du modèle.	Réduire la redondance et minimiser le nombre d'attributs sans évaluations de modèles fréquentes.	Moins coûteux en calculs grâce à une évaluation moins fréquente.	Moins précis que <i>W-DMEA-FS</i> , dépend de la qualité de la fonction de filtrage.
W-QEISS	[Kar+15]	<i>Quasi-Equally Informative Subsets Selection</i> de type <i>wrapper</i> .	Réduire la redondance tout en maximisant la précision de classification.	Meilleure précision en minimisant la redondance et en maximisant la pertinence.	Coût de calcul élevé et peut converger lentement sur de grands ensembles de données.
F-QEISS	[Kar+15]	Variante de <i>QEISS</i> utilisant une méthode de sélection <i>filter</i> pour simplifier la sélection des attributs avec des critères non basés sur un modèle.	Réduire la redondance et augmenter la pertinence des attributs avec un calcul réduit.	Calcul plus rapide et efficace sur de grandes dimensions.	Moins de précision que <i>W-QEISS</i> , moins robuste pour la sélection d'attributs complexes.
W-MOSS	[Ham+07]	Multi-Objective Subset Selection utilisant une méthode de sélection <i>wrapper</i> .	Optimiser le nombre d'attributs sélectionnés et la précision de classification (deux objectifs).	Simple à implémenter et efficace pour des problèmes à deux objectifs.	Moins efficace pour des problèmes à objectifs multiples complexes ; coût de calcul élevé.
F-MOSS	[Ham+07]	Variante de <i>MOSS</i> utilisant une méthode de sélection <i>filter</i> pour une évaluation simplifiée des sous-ensembles.	Réduire la redondance des attributs et optimiser leur pertinence sans évaluation directe de modèle.	Rapide et adapté aux grands ensembles de données avec moins de calculs.	Moins précis que <i>W-MOSS</i> , dépend de la qualité des critères de filtrage.
mRMR	[PLD05]	<i>minimum Redundancy Maximum Relevance</i> pour la sélection d'attributs (de type <i>filter</i>).	Minimiser la redondance et maximiser la pertinence par rapport à la variable cible.	Rapide et efficace pour des problèmes de petite à moyenne dimension.	Ne prend pas en compte la corrélation entre attributs, ce qui peut limiter la performance.
mRMR-BA	[Alo+17]	<i>mRMR</i> suivi d'un <i>Bat Algorithm (BA)</i> pour l'optimisation des sous-ensembles. (Sélection <i>Filter-Wrapper</i>)	1- filtrage <i>mRMR</i> ; 2- optimisation par <i>BA</i> .	Bonne combinaison de filtrage rapide et d'optimisation fine des sous-ensembles.	Peut être complexe à paramétrer ; dépend fortement des paramètres de l'algorithme <i>Bat</i> .
FSPSOTC	[AK17]	Sélection d'attributs avec optimisation par essaim de particules (<i>Wrapper-PSO</i>), spécialement pour le clustering de textes.	Créer de nouveaux sous-ensembles d'attributs informatifs pour le clustering.	Performant pour le clustering de textes et pour éviter les minima locaux.	Spécifique aux problèmes de texte, moins performant sur des ensembles de données non textuelles.

TABLE 2.2 – Comparaison de méthodes de type métaheuristique de sélection d'attributs multi-objectifs avec leurs avantages et inconvénients

À la lecture du Tableau 2.2, nous pouvons en conclure que ces méthodes de recherche métaheuristique, principalement celles utilisant des méthodes de sélection de type *wrapper*, ont démontré leur efficacité pour gérer la sélection d'attributs dans des problèmes complexes de classification¹.

1. Note : Les méthodes de type *filter*, bien que computationnellement plus légères, sont souvent utilisées en combinaison pour un prétraitement rapide et efficace des données.

2.1.2.4 Autres méthodes qui pourraient être considérées.

Bien que les méthodes comparées ci-dessus offrent un aperçu des approches métaheuristiques en sélection d'attributs multi-objectifs, d'autres techniques pertinentes [Jia+23], pourraient également mériter d'être explorées.

Les approches basées sur l'optimisation par essaims de particules (*PSO*) [VPM20], notamment avec des mécanismes adaptatifs ou niching [Guo+21], auraient pu être utilisées pour améliorer la diversité des solutions dans l'ensemble Pareto tout en évitant une convergence prématurée.

De même, les algorithmes de colonies de fourmis (*ACO*) [KN13] présentaient un potentiel intéressant pour capturer les interactions complexes entre attributs grâce à leur représentation en graphe.

Pourraient également être considérés : les algorithmes basés sur la colonie d'abeilles artificielles (*ABC*) [Han22] qui ont démontré leur performance par rapport à d'autres variantes récentes d'algorithmes évolutionnaires, ou encore les approches basées sur la recherche taboue (*Tabu Search*) [Ben+23].

Enfin, des techniques hybrides combinant des approches supervisées et non supervisées, comme l'intégration de *k-means* avec des algorithmes *wrapper* ou *filter-wrapper* (p.ex., *GAK-MEANS* [BAA23]) ont su récemment améliorer la précision de classification tout en réduisant la dimensionnalité des données.

2.2 Positionnement de nos contributions

Les méthodes de sélection d'attributs multi-objectifs se distinguent par leur capacité à optimiser simultanément plusieurs critères souvent contradictoires. C'est dans cet esprit que nous avons développé *GAwC* (*Genetic Algorithm with Multi-Objective Compass*), dont l'objectif est de minimiser le nombre d'attributs tout en maximisant l'exactitude (*accuracy*) et l'*AUC* (*Area Under the Curve*) pour la classification binaire. Ce type d'approche est particulièrement pertinent dans des domaines où l'interprétabilité des résultats est cruciale, tels que la médecine, où les décisions doivent être basées sur des attributs clairs et explicables.

Cependant, ces méthodes de sélection d'attributs présentent des problématiques similaires à ceux rencontrés dans les systèmes de recommandation multicritères. Parmi ces problématiques figurent la gestion de la scalarisation des objectifs, la hiérarchisation de ces derniers, et la détermination d'un compromis adapté aux différents besoins d'optimisation.

Les recherches que nous avons menées autour de *GAwC* reposent sur ce contexte d'optimisation multicritère, en se concentrant spécifiquement sur les données d'entraînement ayant une forte dimensionnalité (c.-à-d., ayant de nombreux attributs ou variables). Ces contraintes de dimensionnalité posent des exigences importantes en termes de robustesse et de performance des algorithmes. Pour répondre à ces exigences, notre méthode *GAwC* utilise un algorithme génétique couplé à une méthode de calcul de fitness multi-objectifs – le *Compass* – permettant de sélectionner dynamiquement les sous-ensembles d'attributs optimaux. Notre al-

gorithme est conçu pour équilibrer plusieurs objectifs afin de se constituer comme une solution robuste et adaptable à plusieurs cas du monde réel.

Les résultats de nos recherches sur *GAWC* ont conduit à une contribution significative, publiée dans le journal Artificial Intelligence in Medicine (Elsevier) en 2022 (revue de premier rang (Q1) en intelligence artificielle selon le Scimago Journal & Country Rank).

Nous présentons la méthode *GAWC* à la section suivante.

2.3 Genetic Algorithm with Multi-Objective Compass - *GAWC*

Dans cette section, nous formalisons tout d'abord le problème multi-objectif. Dans un second temps, nous décrivons la méthode *GAWC* en détaillant :

- l'algorithme génétique original sur lequel elle repose ;
- sa genèse et son concept général ;
- son fonctionnement détaillé.

Enfin, nous réalisons une synthèse des résultats obtenus et nous les discutons.

Cette section utilise des éléments présentés dans [Gut+22].

2.3.1 Formalisation du problème multi-objectifs

Dans notre cas d'études, nous cherchons à sélectionner un faible nombre d'attributs à partir d'un ensemble de données tout en maintenant des niveaux élevés d'*accuracy* et d'*AUC* pour un classifieur donné. Ceci peut être formulé comme un problème multi-objectif. Nous définissons ainsi formellement les fonctions objectifs de notre problème comme suit :

Fonctions objectifs. Soit \mathcal{V} l'ensemble des attributs candidats et \mathcal{S} un sous-ensemble de \mathcal{V} . La cardinalité $|\mathcal{S}|$ est le nombre d'attributs sélectionnés dans \mathcal{V} et représente le premier objectif à minimiser. Cette fonction est définie comme suit :

$$f_1(\mathcal{S}) = |\mathcal{S}| \quad (2.1)$$

Le deuxième objectif à maximiser est l'exactitude (*accuracy*) de la classification binaire de l'algorithme d'apprentissage. Cette fonction est définie comme suit :

$$f_2(\mathcal{S}) = \frac{vp + vn}{vp + vn + fp + fn} \quad (2.2)$$

où vp , vn , fp , fn désignent respectivement les vrais positifs, les vrais négatifs, les faux positifs et les faux négatifs.

Enfin, le dernier objectif que nous considérons est l'*AUC*. Un tel critère est considéré comme l'un des plus importants à prendre en compte dans le domaine de la médecine, car comme le coefficient de corrélation de Matthews, il représente l'une des meilleures métriques pour observer la qualité de la classification [CTJ21]. En plus de l'*accuracy*, l'*AUC* résume les

quatre métriques de toute la matrice de confusion : tp (vrais positifs), tn (vrais négatifs), fp (faux positifs) et fn (faux négatifs). Ainsi, plus la valeur de l' AUC est élevée, meilleur est le classifieur en termes de qualité. Des exemples de courbes ROC et leurs AUC associées, allant d'un classifieur parfait ($AUC = 1$) à un classifieur basé sur un choix aléatoire ($AUC = 0.5$), sont donnés à la Figure 2.1.

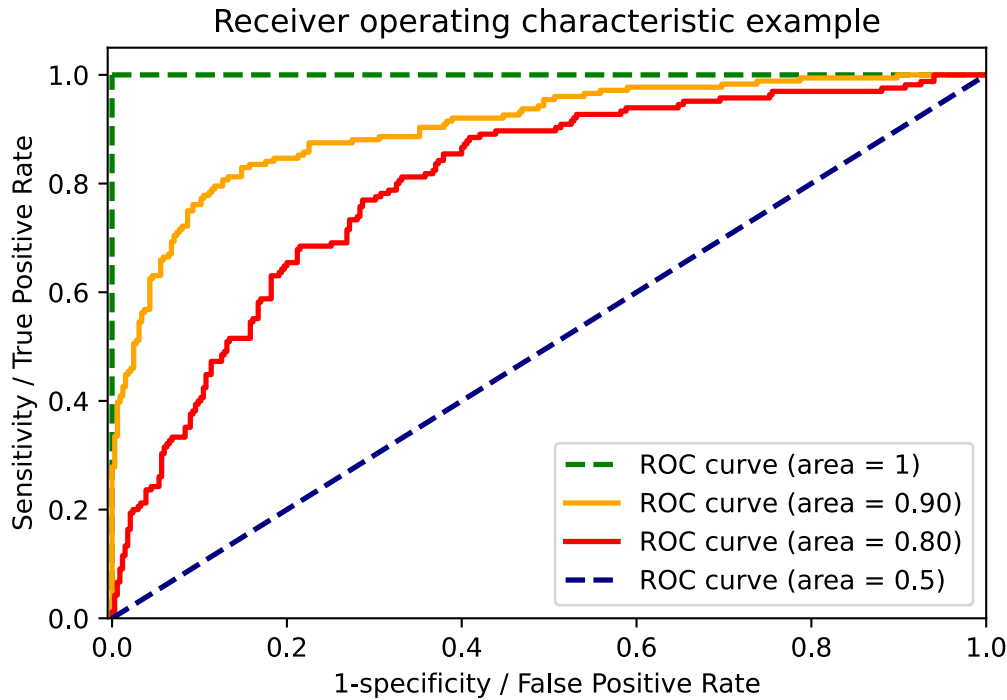


FIGURE 2.1 – Courbes ROC et valeurs d' AUC .

Ce dernier objectif à maximiser peut être défini plus précisément comme suit :

$$f_3(\mathcal{S}) = \int_0^1 ROC(x) dx \quad (2.3)$$

où $ROC(x)$ désigne la courbe ROC (*Receiver Operating Characteristic*).

2.3.2 La méthode GAwC

2.3.2.1 Algorithme génétique original

L'algorithme génétique (voir Algorithme 4) que nous utilisons dans $GAwC$ repose sur [Hol92]. Il repose sur le principe de sélection naturelle, et vise à trouver une solution optimale au sein d'un espace de recherche défini. Il démarre avec la création aléatoire d'une population d'individus, chaque individu représentant une solution potentielle. Chaque individu est évalué à

l'aide d'une fonction de fitness, qui mesure son adaptation à la tâche. La sélection est effectuée par un tournoi où les meilleurs individus sont choisis pour se reproduire. Un croisement à deux points est appliqué, générant une nouvelle génération de descendants en combinant les caractéristiques des parents. Nous avons choisi le croisement à deux points qui reste un bon compromis entre diversité et exploitation, bien que d'autres solutions pourraient être envisagées comme le croisement uniforme, le croisement basé sur la fitness (*adaptive crossover*) ou encore le croisement guidé par heuristique. Ensuite, une étape de mutation intervient avec une probabilité donnée, permettant de maintenir la diversité génétique de la population, bien que le meilleur individu soit protégé de cette mutation. La population des parents et des descendants fusionne, formant une nouvelle génération qui est évaluée. À chaque génération, le meilleur individu est sauvegardé, et s'il surpasse les précédents, il devient le nouvel individu optimal. Ce processus est répété pour un nombre fixe de générations, et l'individu le plus adapté rencontré au cours de toutes les générations est retourné en tant que meilleure solution trouvée par l'algorithme.

Algorithme 4 : Principe général de l'algorithme génétique employé [Hol92]

Entrée : n : Nombre d'individus.

individualSize : Taille de chaque individu.

nbGens : Nombre de générations (itérations).

mutationRate : Probabilité de muter un bit d'un individu.

Sortie : Meilleur individu de la population.

- 1 **Initialisation** : Créer une population aléatoire de n individus.
 - 2 **Évaluer** chaque individu de la population selon la fonction de fitness.
 - 3 **Sauvegarder** dans $xMax$ l'individu le plus adapté de la population.
 - 4 **for** $i \leftarrow 1$ **to** $nbGens$ **do**
 - 5 **Sélection par tournoi** : la population est regroupée aléatoirement par paires. La fitness des deux parents de chaque paire est comparée et seul le parent avec la meilleure fitness est conservé.
 - 6 **Croisement** pour construire la génération des descendants : un opérateur de croisement à deux points est utilisé.
 - 7 **Évaluer** chaque enfant, selon la fonction de fitness.
 - 8 **Fusionner** les parents ($n/2$ individus) et les enfants ($n/2$ individus) pour retrouver une population de n individus.
 - 9 **Muter** chaque bit de la population avec une probabilité égale à $mutationRate$ (les bits du meilleur individu ne sont pas mutés) – *bit flip* –
 - 10 **Évaluer** chaque individu muté, selon la fonction de fitness.
 - 11 **Sauvegarder** dans $xMaxGen$ l'individu le plus adapté de la population de la génération courante.
 - 12 **if** $xMax < xMaxGen$ **then**
 - 13 $xMax = xMaxGen$
 - 14 **Retourner** $xMax$, l'individu le plus adapté rencontré au cours de toutes les générations.
-

Des compléments d'information sur les détails spécifiques de configuration de l'Algorithme 4 utilisé dans la méthode *GAwC* sont données dans l'Annexe A.3.

2.3.2.2 Genèse et concept de *GAwC*

La méthode de *Compass* s'étant montrée efficace dans la suite des méthodes *Gorthaur* (voir Chapitre 1) que nous avons conçus en bi-objectifs, nous avons eu l'intuition de l'adapter et de la décliner en multi-objectif (3D) dans ce nouveau cadre applicatif. Ainsi, *GAwC* (*Genetic Algorithm with Multi-Objective Compass*) [Gut+22], repose sur une approche hybride combinant un algorithme génétique (Voir Algorithme 4) avec une méthode de calcul de fitness employant la méthode *Compass* [MS08]. Cette fois la méthode est étendue pour pouvoir optimiser trois objectifs dans le cadre de la sélection d'attributs pour des problèmes de classification binaire en médecine. Notre méthode vise ainsi à sélectionner un sous-ensemble minimal d'attributs pertinents tout en maintenant une bonne qualité et une bonne *accuracy* de classification.

De plus, l'utilisation de la méthode *Compass* [MS08] pour la définition et l'agrégation des fonctions de fitness rapproche notre méthode du principe de fonctionnement de *MOEA/D* (Multi-Objective Evolutionary Algorithm based on Decomposition) [ZL07]. *MOEA/D* divise un problème multi-objectif en plusieurs sous-problèmes scalaires à l'aide de méthodes de décomposition, où chaque sous-problème est optimisé indépendamment mais de manière coopérative. De manière analogue, le *Compass* agit comme un mécanisme de décomposition, en exploitant une agrégation multicritère qui permet d'équilibrer efficacement les objectifs contradictoires.

2.3.2.3 *GAwC* : Genetic Algorithm with Multi-Objective Compass

Dans cette sous-section, nous présentons le détail de la méthode de sélection d'attributs multicritère - *GAwC* :

1. Dans un premier temps, nous expliquons la méthode *Compass* 3D, et nous décrivons la manière dont elle est utilisée pour calculer le score multicritères d'une solution candidate ;
2. Nous expliquons ensuite le calcul des scores à partir des fonctions objectifs et leur utilisation par la méthode *Compass* ;
3. Enfin, nous détaillons l'algorithme *GAwC* et nous montrons comment le score multicritères de chaque solution candidate peut être pris en compte par l'algorithme génétique employé (Voir Algorithme 4) pour résoudre le problème multi-objectifs.

Le *Compass* à 3 Dimensions. Le *Compass* à trois dimensions s'inspire des méthodes *Compass* à deux dimensions [MS08] et donc de *Gorthaur* [Gut+19b ; Gut+21 ; Let+24]. Il vise à agréger les résultats de trois fonctions de score (sc_1, sc_2, sc_3) avec des valeurs comprises entre $[0, 1]$ en un seul score multicritère.

Le score multicritère S est obtenu en projetant le vecteur de score $\vec{d} = (sc_1, sc_2, sc_3)$ sur un vecteur de référence \vec{c} représentant un compromis cible entre les trois scores (voir figure 2.2).

Nous avons ainsi : $S = \|\vec{d}\| \cos \alpha$ avec $\cos \alpha = \frac{\vec{d} \cdot \vec{c}}{\|\vec{d}\| \|\vec{c}\|}$ et donc :

$$S = \frac{\vec{d} \cdot \vec{c}}{\|\vec{c}\|}$$

Comme détaillé dans [Gut+21], ce score agrégé peut être utilisé pour évaluer les solutions candidates d'un problème d'optimisation multicritère en fonction d'un compromis donné.

Avant de les utiliser dans un *Compass 3D*, les fonctions objectives du problème d'optimisation doivent être normalisées de manière à être uniformément distribuées entre 0 et 1.

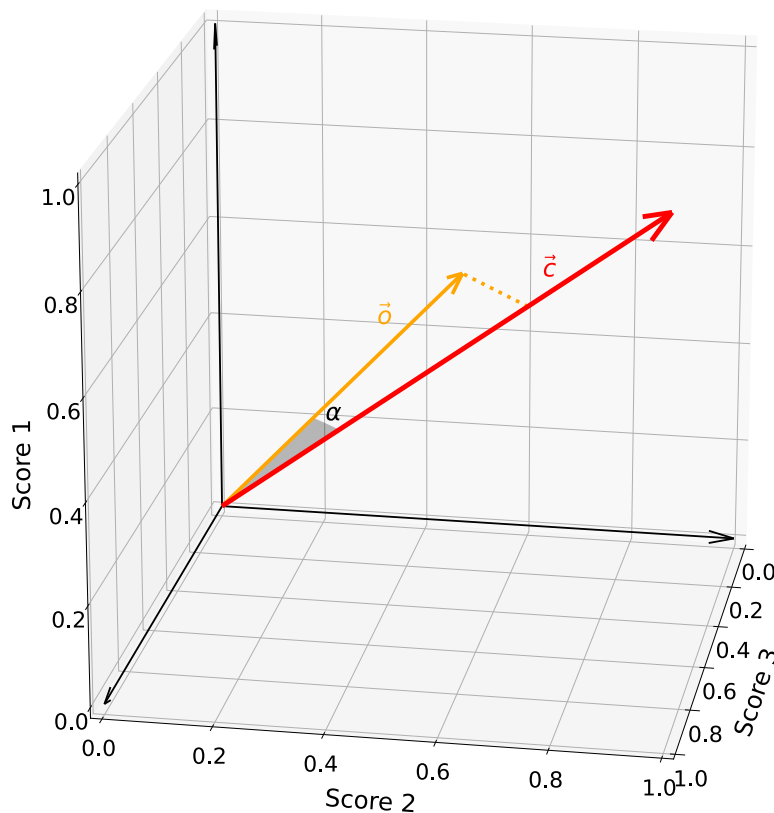


FIGURE 2.2 – Projection d'un vecteur de score sur le vecteur de référence avec le *Compass 3D*.

Note sur l'agrégation de plusieurs critères. L'agrégation de plusieurs critères pour former une mesure synthétique est une problématique classique en Recherche Opérationnelle (RO), particulièrement dans les domaines du multi-objectif et de la prise de décision multicritère. Ce cadre théorique repose souvent sur des approches telles que les fonctions d'utilité additive ou multiplicative [Kee93], les méthodes d'optimisation pondérée [CH08], ou encore les processus

d'analyse hiérarchique (*AHP, Analytic Hierarchy Process*) [CL01 ; Kar82]. Ces méthodes permettent de combiner plusieurs dimensions hétérogènes dans un cadre unifié, tout en tenant compte des priorités ou des poids relatifs associés à chaque critère.

Dans le cadre de *GAWC*, l'objectif est de fournir une synthèse efficace des scores en maintenant une interprétabilité. En outre, la normalisation des scores dans l'intervalle $[0, 1]$ facilite l'agrégation et l'analyse comparative, tout en respectant des conditions de monotonie et d'équidistance souvent requises dans de telles méthodes.

Fonctions objectifs et scores relatifs. La fonction de score utilisée pour la maximisation de la fonction objectif $f_1(\mathcal{S})$ (voir Équation 2.1) pendant le processus de recherche est le "score du nombre d'attributs" (Sc_{Feat}) défini comme

$$Sc_{Feat} = 1 - \frac{\rho}{|\mathcal{S}|} \quad (2.4)$$

où ρ est un facteur appliqué aux ensembles de données de haute dimension visant à converger plus rapidement vers une solution avec un faible nombre d'attributs. Par exemple, [Kar+15] fixe le nombre minimum d'attributs à 20. Ainsi, ici, ρ peut être considéré à la fois comme un moyen de fixer le nombre minimum d'attributs à sélectionner et comme une sorte de taux d'apprentissage garantissant que la convergence vers une solution optimale acceptable ne nécessite pas trop de générations.

La fonction de score utilisée pour la maximisation de la fonction objectif $f_2(\mathcal{S})$ (voir Équation 2.2) pendant le processus de recherche est le "score d'exactitude" (*accuracy : Acc*) Sc_{Acc} . Puisque l'*accuracy* varie de 0 à 1, elle peut être directement utilisée comme une fonction de score. Ainsi,

$$Sc_{Acc} = f_2(\mathcal{S}) \quad (2.5)$$

Concernant la troisième fonction objectif (voir Équation 2.3), nous considérerons que l'*AUC* varie de 0,5 à 1 du fait que l'*AUC* d'une classification aléatoire est de 0,5. Par conséquent, un score pertinent pour l'*AUC* est une valeur normalisée de $f_3(\mathcal{S})$ sur une échelle de 0 à 1. Nous appelons ce score le "score de l'*AUC*" (Sc_{AUC}) afin de le différencier de la valeur réelle de l'*AUC* (score brut).

Le "score de l'*AUC*" (Sc_{AUC}) peut être défini par :

$$Sc_{AUC} = (f_3(\mathcal{S}) - 0.5) \times 2 \quad (2.6)$$

L'algorithme *GAWC*. L'algorithme *GAWC* est une implémentation de l'algorithme génétique (Voir Algorithme 4) inspirée par [Hol92] et [Mez+19] et présentée précédemment. Il vise à trouver une solution optimale de Pareto pour le problème à trois objectifs consistant à sélectionner un sous-ensemble minimal d'attributs à fournir à un classifieur tout en maximisant à la fois l'*accuracy* et le score *AUC-ROC* du classifieur. Comme présenté précédemment, la recherche de la solution dans l'espace de recherche est guidée par un vecteur de référence qui indique un compromis entre les trois objectifs avec lequel la solution optimale doit s'aligner.

Les individus de l'algorithme *GAwC* sont représentés par un vecteur binaire de taille égale au nombre total d'attributs (par exemple, 48 pour le jeu de données fourni par le CHU d'Angers intitulé : ASA Drug Intake). Les bits dans le vecteur indiquent la présence ou l'absence de chaque caractéristique dans l'individu correspondant. La fitness d'un individu est calculée en utilisant la méthode du *Compass* à trois dimensions présentée précédemment.

Plus précisément, pour chaque génération $g_i (i \in [1, nbGens])$ générée par *GAwC* et chaque individu $x_{j,g_i} (j \in [1, n])$ dans la population de la génération, l'algorithme calcule² le score d'*accuracy* de l'individu $Sc_{Acc}(x_{j,g_i})$ (voir Équation 2.5), son score du nombre d'attributs $Sc_{Feat}(x_{j,g_i})$ (voir Équation 2.4) et son score de *AUC* $Sc_{AUC}(x_{j,g_i})$ (voir Équation 2.6).

La capacité de l'individu à être à la fois précis, avec un faible nombre d'attributs, et une bonne qualité de classification (*AUC*) est alors exprimée par le vecteur de score $\vec{o}_{x_{j,g_i}}$ tel que :

$$\vec{o}_{x_{j,g_i}} = (Sc_{Feat}(x_{j,g_i}), Sc_{Acc}(x_{j,g_i}), Sc_{AUC}(x_{j,g_i}))$$

Comme présenté précédemment, la fitness de l'individu est obtenue en projetant son vecteur de score $\vec{o}_{x_{j,g_i}}$ sur le vecteur de référence prédéfini \vec{c} et peut être calculée comme suit :

$$f(x_{j,g_i}) = |\vec{o}_{x_{j,g_i}}| \cos \alpha_{x_{j,g_i}} \quad (2.7)$$

où, comme présenté dans la figure 2.2, α_{j,g_i} est l'angle entre le vecteur de référence \vec{c} et le vecteur de score de l'individu $\vec{o}_{x_{j,g_i}}$.

À chaque génération g_i générée par *GAwC*, l'algorithme détermine $x_{max}(g_i)$, l'individu le plus adapté de la population de la génération g_i , tel que :

$$x_{max}(g_i) = \underset{j \in [1, n]}{f(x_{j,g_i})}$$

Enfin, si nécessaire, $x_{max}(g_i)$ est utilisé pour mettre à jour x_{max} , l'individu le plus adapté rencontré au cours de toutes les générations :

$$x_{max} = \underset{i \in [1, nbGens]}{\max} x_{max}(g_i)$$

Ainsi, après que les $nbGens$ générations aient été générées par l'algorithme génétique, x_{max} est l'individu le plus adapté rencontré au cours de toutes les générations.

La figure 2.3 illustre l'évolution du vecteur de score de x_{max} au fil des générations i , $i + 50$ et $i + 100$ et montre la fitness de l'individu correspondant obtenue en projetant le vecteur de score sur le vecteur de référence \vec{c} .

Dans notre implémentation de l'algorithme génétique et selon le protocole suivi par [Li+20], nous fixons $nbGens$, le nombre de générations, à 100.

Un exemple concret et réel sur le jeu de données *Musk1*³ est disponible à l'Annexe A.4 avec quatre vues du *Compass* pour mieux prendre en compte l'évolution au cours du temps.

2. lorsqu'on utilise les attributs représentés par l'individu

3. Musk1 - [https://archive.ics.uci.edu/ml/datasets/Musk+\(Version+1\)](https://archive.ics.uci.edu/ml/datasets/Musk+(Version+1))

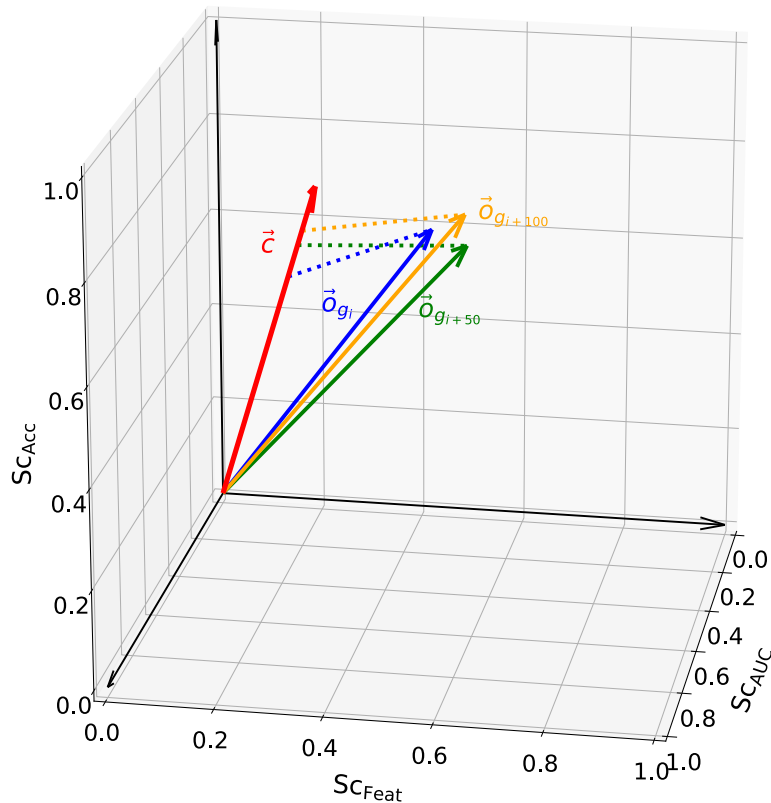


FIGURE 2.3 – Vecteurs de score et fitness des meilleurs individus aux générations g_i , g_{i+50} et g_{i+100}

En résumé, *GAWC* fonctionne comme l'algorithme général défini dans la section 4 mais incorpore le *Compass* 3D comme fonction de fitness dans son processus d'évaluation. En d'autres termes, *GAWC* évalue chaque individu de la population, chaque descendant et chaque individu muté (voir lignes 2, 7, 10 de l'algorithme défini dans la section 4, respectivement), en fonction de la fitness f_{g_i} qui est calculée comme exprimé dans l'Équation 2.7.

2.3.3 Synthèse des résultats et discussion sur la méthode *GAWC*

Cette sous-section est une synthèse des résultats obtenus. Le détail de ces résultats est disponible dans l'article [Gut+22].

Résultats empiriques. L'algorithme *GAWC* (Genetic Algorithm with multi-objective Compass) a été évalué sur six jeux de données réelles (principalement dans le domaine médical, voir Tableau 2.3 et [Gut+22]) et comparé à plusieurs méthodes concurrentes de sélection d'attributs multi-objectif (voir Tableau 2.2 : *W-DMEA-FS*, *F-DMEA-FS*, *W-QEISS*, *F-QEISS* et *W-MOSS*).

Les résultats montrent que *GAWC* surpasse toutes les autres méthodes en termes de :

Nom du jeu de données	Nombre d'instances	Nombre d'attributs	Type catégoriel	Type continu	Source du jeu de données
Breast Cancer	569	30	0	30	UCI ML Repository
Cardiotocography	2126	21	0	21	UCI ML Repository
Diabetes	768	8	8	0	UCI ML Repository et Kaggle
Heart	270	13	7	6	UCI ML Repository
Musk1	476	166	0	166	UCI ML Repository
ASA-DI	822	48	30	18	Hôpital Universitaire d'Angers

TABLE 2.3 – Jeux de données utilisés.

- **exactitude (*accuracy*)** : *GAwC* a obtenu des résultats supérieurs ou équivalents en termes d'*accuracy* par rapport aux méthodes concurrentes telles que *W-DMEA-FS*, *F-DMEA-FS*, *W-QEISS*, *F-QEISS* et *W-MOSS* pour tous les jeux de données.
- **nombre d'attributs sélectionnés** : *GAwC* a également montré une performance robuste en minimisant le nombre d'attributs nécessaires tout en maintenant une haute *accuracy*, particulièrement sur des ensembles de données de moyenne et haute dimension.

Plus précisément si nous analysons les résultats obtenus sous trois angles (le meilleur : BEST, le pire : WORST, et la moyenne : AVERAGE) :

- **BEST** : *GAwC* a surpassé les autres méthodes pour la *accuracy* maximale atteinte sur chaque ensemble de données tout en maintenant un nombre d'attributs sélectionnés faible et optimal.
- **WORST** : Même pour les cas les plus défavorables, *GAwC* a maintenu une meilleure *accuracy* et un nombre d'attributs sélectionnés compétitif.
- **AVERAGE** : En moyenne, *GAwC* a obtenu de meilleures performances en termes d'*accuracy* et a sélectionné un nombre d'attributs plus faible par rapport aux méthodes concurrentes.

De plus, les attributs sélectionnés par *GAwC* pour le jeu de données ASA Drug Intake (ASA-DI) fournis par le CHU d'Angers, ont été validés comme cliniquement pertinents par le Pr Pierre Abraham (expert clinicien), renforçant l'utilité pratique de la méthode.

Synthèse et discussion. La synthèse des résultats présentés ci-dessus, démontre que *GAwC* est une méthode prometteuse pour les problèmes de sélection d'attributs multicritère dans des applications médicales, en garantissant à la fois la qualité de la classification (via l'*AUC* et l'*accuracy*) et l'efficacité en réduisant le nombre de variables. De plus, *GAwC* se distingue des autres méthodes de sélection d'attributs sur plusieurs aspects importants, entre autres, dans le domaine médical :

- L'inclusion de l'*AUC* comme critère d'optimisation contrairement aux approches traditionnelles qui se concentrent principalement sur l'*accuracy* ou le nombre d'attributs sélectionnés. Notre méthode considère ainsi l'aire sous la courbe ROC (*AUC*) comme un

critère clé. L'*AUC* est particulièrement pertinente dans le domaine médical car elle permet de mieux distinguer entre les classes de patients sains et malades, en prenant en compte à la fois la sensibilité et la spécificité. Cela permet de garantir la qualité de la classification pour les applications où une détection précise des profils de patients est essentielle.

- L'utilisation d'un *Compass* multi-objectif (3D) pour guider la recherche et ajuster le compromis entre les trois objectifs (minimisation du nombre d'attributs, maximisation de l'*accuracy* et maximisation de l'*AUC*). Ce mécanisme de fitness simplifie la paramétrisation des objectifs en définissant un simple vecteur de référence 3D, rendant la direction de recherche plus explicite et visuelle pour le praticien (utilisateur de la méthode).
- La méthode *Compass* agrège ces objectifs en un seul objectif composite à optimiser, ce qui permet de simplifier le processus d'optimisation multi-objectifs tout en offrant une grande flexibilité de paramétrage (via l'angle α). Cependant, cela peut entraîner une perte de la spécificité de chaque objectif individuel, car la méthode vise à obtenir un compromis général pour l'ensemble des objectifs.

Un point clé à prendre en compte concernant l'intégration d'une méthode *Compass* est qu'elle fonctionne de manière optimale avec des objectifs indépendants ou non corrélés. Cependant, elle présente une limitation notable : elle a tendance à lisser progressivement la solution optimale de chaque objectif individuel en cherchant à agréger une solution de compromis pour l'ensemble des objectifs. En revanche, cette méthode est facilement paramétrable et permet de définir les degrés d'importance accordés à chaque objectif.

En résumé : *GAwC* doit être considérée comme une méthode de sélection multicritère agrégée en un seul objectif, car elle utilise une fonction de fitness combinée qui représente plusieurs objectifs avec un vecteur de référence dans un espace à trois dimensions. Contrairement aux méthodes multi-objectifs classiques, *GAwC* n'utilise pas de tri non dominé ou de distance d'encombrement comme par exemple dans *NSGA-II*, mais maximise plutôt une fonction de fitness unique définie par l'angle α du *Compass*. Cette approche simplifie l'optimisation mais ne permet pas une exploration aussi fine du front de Pareto.

Un des autres avantages de *GAwC* réside dans sa capacité à ajuster dynamiquement un compromis entre les différents objectifs matérialisés par le *Compass*. Ce compromis est calculable ou paramétrable via le paramètre α du *Compass*. Cette flexibilité permet aux utilisateurs de la méthode (dans notre cas des praticiens) de moduler la priorité entre les objectifs en fonction des besoins spécifiques de l'application p.ex., la minimisation du nombre d'attributs et la maximisation de l'*accuracy*. Cependant, bien que l'agrégation des objectifs permette une optimisation plus simplifiée pour un utilisateur non initié, elle peut aussi limiter la capacité de la méthode à atteindre une solution optimale plus spécifique à chaque objectif. Cette limitation doit être prise en considération lorsqu'on souhaite mettre en application *GAwC* à des problèmes où l'équilibre entre les objectifs reste critique.

Pour conclure, il sera important de retenir que *GAwC* est principalement une méthode *wrapper* qui se distingue par son approche d'agrégation multicritère et sa flexibilité/simplicité de

paramétrage, rendant la méthode particulièrement utile pour les problèmes où un compromis contrôlé entre plusieurs objectifs est nécessaire.

2.3.4 Réflexion critique

2.3.4.1 Rapprochement avec MOEA/D

Les algorithmes *Genetic Algorithm with Multi-objective Compass* (GAwC) et *MOEA/D* (Multi-Objective Evolutionary Algorithm based on Decomposition) présentent plusieurs similarités et différences dans leur approche de l'optimisation multi-objectifs.

Similitudes :

- **Approche basée sur la décomposition :**
 - *GAwC* utilise un *compass 3D* pour guider la recherche en projetant chaque solution sur une référence vectorielle définissant les priorités entre les objectifs. Cela peut être vu comme une forme de décomposition.
 - De son côté, *MOEA/D* décompose le problème multi-objectifs en plusieurs sous-problèmes scalaires, chacun correspondant à une combinaison pondérée des objectifs.
- **Optimisation :** Les deux algorithmes cherchent à optimiser simultanément plusieurs objectifs conflictuels. Par exemple, *GAwC* vise à minimiser le nombre d'attributs tout en maximisant l'exactitude et l'*AUC*, tandis que *MOEA/D* optimise des objectifs multiples dans un cadre général.
- **Utilisation des concepts évolutionnaires :** Les deux méthodes reposent sur des mécanismes évolutionnaires tels que la sélection, le croisement et la mutation pour explorer efficacement l'espace de recherche.
- **Solutions sur le front de Pareto :** Les deux approches produisent un ensemble de solutions non dominées représentant un front de Pareto optimal, permettant aux utilisateurs de choisir le compromis le plus adapté à leurs besoins.

Différences :

- **Métriques d'évaluation :**
 - *GAwC* inclut l'*AUC* (*Area Under the ROC Curve*) comme critère principal d'optimisation, ce qui est particulièrement pertinent dans les applications médicales.
 - *MOEA/D*, en revanche, n'intègre pas directement des métriques spécifiques aux problèmes, mais s'appuie sur une formulation généraliste.
- **Guidage de la recherche :**
 - *GAwC* utilise un *compass 3D* basé sur un vecteur de référence unique, simplifiant la paramétrisation et rendant la direction de recherche explicite.

- *MOEA/D* divise l'espace d'objectifs à l'aide de poids pour guider la recherche, ce qui requiert une configuration qu'un utilisateur pourrait considérer comme potentiellement plus complexe.
- **Contextes d'application :**
 - *GAwC* est spécifiquement conçu pour la sélection d'attributs dans les problèmes de classification médicale, en s'orientant implicitement vers des métriques comme la sensibilité et la spécificité.
 - *MOEA/D* est plus généraliste et peut être appliqué à une large gamme de problèmes multi-objectifs sans ajustements spécifiques.
- **Complexité algorithmique :**
 - Grâce au *compass 3D*, *GAwC* offre une approche simplifiée pour ajuster les priorités entre objectifs.
 - *MOEA/D* nécessite un réglage potentiellement plus minutieux des poids de décomposition pour une performance optimale, ce qui peut augmenter la complexité auprès d'utilisateurs n'étant pas du métier.

Conclusion : Bien que *GAwC* et *MOEA/D* partagent des principes communs, *GAwC* se distingue par son adaptation spécifique aux applications médicales et son mécanisme de guidage potentiellement plus simplifié grâce au *compass 3D*. En revanche, *MOEA/D* reste une méthode plus générale, adaptée à une grande variété de problèmes d'optimisation multi-objectifs.

2.3.4.2 Éthique des décisions algorithmiques.

Les méthodes de sélection d'attributs qu'elles soient multi-objectifs ou multicritères, telles que *GAwC* (*Genetic Algorithm with Multi-Objective Compass*), s'inscrivent dans une réflexion sur l'éthique des décisions algorithmiques.

Le paramétrage des objectifs de *GAwC* via un vecteur de référence en trois dimensions (c.-à-d., le *compass*) peut ainsi toucher aux débats éthiques sur la transparence et la responsabilité dans l'intelligence artificielle. En effet, le vecteur de référence dans *GAwC* est suffisamment explicite pour permettre aux utilisateurs de comprendre et de contrôler les compromis effectués par l'algorithme. Cela se rapporte ainsi à la demande d'*explicabilité* en IA, où les parties prenantes (dans notre cas, les praticiens hospitaliers) doivent être en mesure de comprendre le processus de prise de décision des systèmes d'IA pour garantir la confiance et la conformité éthique.

Ainsi, l'approche de *GAwC* en matière d'optimisation multicritère pour la sélection d'attributs médicaux met non seulement en lumière l'interaction entre la complexité et la simplicité mais aussi les impératifs éthiques et de transparence dans la prise de décision algorithmique. *GAwC* ne se contente pas d'optimiser des critères mathématiques ; il engage également des questions sur la nature de la prise de décision dans des domaines complexes et sensibles comme la médecine.

DEUXIÈME PARTIE

Apprentissage profond et données complexes pour l'estimation et la génération

L'apprentissage profond repose sur la capacité des réseaux neuronaux à traiter de grandes quantités de données plus ou moins complexes et à en extraire des représentations pertinentes avec ou sans supervision explicite selon les cas. Ces modèles sont particulièrement adaptés pour modéliser des relations complexes entre les entrées (données) et les sorties (résultats souhaités) à travers plusieurs couches non linéaires. Les réseaux apprennent de manière hiérarchique, c'est-à-dire qu'ils construisent des représentations abstraites de plus en plus complexes au fur et à mesure que les informations traversent les différentes couches du réseau [LBH15].

Nos travaux de recherche ont porté sur l'apprentissage profond appliqué à la génération de musique symbolique et sur l'apprentissage profond guidé par la physique appliqué à l'évaluation de systèmes dynamiques. Ces deux domaines reposent sur des données complexes, mais la nature de l'apprentissage et les méthodes utilisées diffèrent, car les problématiques liées à la complexité ne sont pas les mêmes. Ainsi, différentes approches en apprentissage profond ont été conçues pour aborder des problèmes variés dans divers contextes, en s'adaptant à la nature et à la quantité des données disponibles [Goo16].

Ci-dessous, nous rappelons quelques spécificités des méthodes d'apprentissage profond qui nous ont particulièrement intéressé dans nos travaux : 1) dans le domaine musical 2) dans l'évaluation des systèmes dynamiques non linéaires, avec une application aux véhicules terrestres de la Direction Générale de l'Armement (DGA).

Apprentissage profond dans le domaine musical. Dans le domaine de la génération de musique symbolique, les réseaux neuronaux profonds sont entraînés sur de vastes ensembles de données constitués de représentations symboliques de la musique, telles que les partitions ou les séquences MIDI [Hua+18]. Ces représentations capturent des éléments essentiels de la musique, tels que les notes, les accords et les rythmes. Toutefois, il est important de noter que, dans ce contexte, les réseaux n'incorporent pas explicitement de connaissances théoriques musicales, comme c'est souvent le cas dans d'autres domaines (par exemple, en physique). Ils apprennent à partir des données elles-mêmes, suivant une approche entièrement basée sur les données (*data-based*).

Afin de préparer les données d'entraînement et les fournir aux modèles d'apprentissage profond sous forme d'une représentation structurée et employable par ces modèles, l'une des techniques couramment employées dans ce domaine est la *tokenisation musicale*. La tokenisation permet de convertir les éléments musicaux en séquences de tokens. Ces séquences sont ensuite utilisées pour former des modèles séquentiels spécifiques à la musique [Hua+18 ; SUV18]. Parmi les méthodes les plus utilisées, on retrouve les *LSTM* [Hoc97] ou les *Transformers* [Vas+17] qui apprennent à modéliser des structures musicales complexes, y compris les relations harmoniques et mélodiques, de manière implicite à partir des données. Les performances de ces modèles dépendent principalement de la qualité et de la quantité des données disponibles, et les représentations apprises restent souvent non interprétables pour un humain.

Ainsi, dans ce cadre entièrement *data-based*, l'amélioration des performances est fortement dépendante de la qualité des données d'entraînement et des techniques d'optimisation de la

représentation, telles que la *tokenisation*. Celles-ci visent à ce que les modèles puissent mieux capturer la structure des données musicales plus ou moins complexes [HY20].

Apprentissage dans les systèmes dynamiques non linéaires. Dans le cadre des systèmes dynamiques non linéaires appliqués aux véhicules terrestres, deux types d'approche peuvent être adaptés : 1) les approches basées sur les données [Liu+19; Rud+17; SL19] (*data-based*), qui exploitent des ensembles de données réels pour identifier, modéliser ou estimer le comportement du système à l'aide de techniques d'apprentissage ; 2) les approches basées sur les modèles [AO03; Esf+23; ZB09] (*model-based*), qui s'appuient sur des équations mathématiques et des lois physiques pour représenter, estimer ou simuler le comportement dynamique du système. Une combinaison de ces deux approches est possible et sera donc considérée comme une approche « hybride » [CLL23]. Notons que pour la partie *data-based* des approches hybrides, l'apprentissage profond est souvent adopté. Dans ce cadre, l'apprentissage combine des données expérimentales avec des modèles physiques bien établis, tout en exploitant les capacités des réseaux neuronaux pour modéliser les relations complexes entre les paramètres du système. Les réseaux neuronaux profonds sont ainsi formés à partir de données mesurées ou simulées à l'aide de modèles physiques [Raj11], tels que ceux sur lesquels nous nous sommes intéressés dans le cadre de la thèse de Fabien Lioni entre autre appliqués à l'évaluation de la dynamique latérale des véhicules.

Ces types d'approches hybrides (*model-based*, *data-based*) tirent partis à la fois des données réelles pour enrichir les simulations et des modèles physiques pour contraindre l'apprentissage [RPK19]. En complément, des techniques telles que l'*ABC-SMC* [Ton+09] peuvent être utilisées pour affiner les estimations de paramètres, mais c'est principalement l'apprentissage profond qui permet de capturer les dynamiques complexes et non linéaires du système dans un contexte possiblement temps-réel. L'utilisation conjointe de modèles physiques et de données permet de réduire la dépendance aux données expérimentales et d'améliorer l'interprétabilité des résultats, tout en bénéficiant des capacités de généralisation des réseaux neuronaux [RPK19].

L'incorporation de modèles physiques permet également de réduire la dépendance aux données expérimentales, en particulier lorsque ces données sont coûteuses ou difficiles à obtenir. De plus, ces modèles offrent une interprétation plus claire des résultats, car les paramètres estimés sont directement liés aux propriétés physiques du système étudié [RPK19].

Comparaison et complémentarité des approches. Les deux approches – entièrement *data-based* pour la musique et *hybride* pour les systèmes dynamiques – offrent des perspectives complémentaires pour l'apprentissage à partir de données complexes.

Dans le cas de la musique, nous avons employé les données comme représentations symboliques sans modèle sous-jacent fondé sur des lois ou des règles explicites musicales. Les performances des modèles dépendent donc entièrement de la disponibilité et du traitement des données. En revanche, dans les systèmes dynamiques, nous avons employé des données

provenant à la fois de mesures expérimentales et de simulations basées sur des modèles physiques, qui servent de guide pour améliorer l'apprentissage et surmonter les limitations liées aux données.

L'amélioration des performances que nous avons obtenue dans le domaine de la musique repose sur l'optimisation des représentations complexes, tandis que dans les systèmes dynamiques, elle est obtenue en combinant des modèles physiques avec les données expérimentales. Cette approche hybride a permis de réduire l'incertitude et d'améliorer la généralisation des modèles dans des environnements complexes [LBH15].

Ces deux types de solution à des problématiques d'apprentissage représentent donc des stratégies complémentaires pour traiter des problèmes complexes dans des environnements riches en données, qu'il s'agisse de données musicales ou de systèmes physiques [BGC17].

Contributions de la Partie II. Le sujet de cette Habilitation à Diriger des Recherches se concentre plus particulièrement sur l'apprentissage automatique pour la génération et l'exploration avec un focus particulier sur l'adaptation et la contextualisation pour données complexes. Cette partie met l'accent sur l'apprentissage profond avec données complexes : 1) pour l'estimation de paramètres en environnement dynamique non linéaire possiblement temps-réel ou encore l'estimation d'état ; 2) pour la génération de musique symbolique. Cette partie présente ainsi des travaux autour de deux cadres applicatifs différents :

1. **l'évaluation de comportements dynamiques non linéaires appliquée entres autres aux véhicules terrestres** (voir Chapitre 3 de cette partie). Ce troisième chapitre cible plus précisément l'estimation de paramètres dans les systèmes dynamiques non linéaires, tels que les dynamiques latérales de véhicules, à travers l'utilisation de méthodes de filtrage statistique et d'apprentissage profond. Il traite également de l'estimation d'état en temps réel par apprentissage profond et estimation à horizon glissant.
2. **l'optimisation des représentations symboliques musicales pour l'apprentissage Profond** (voir Chapitre 4 de cette partie). Le quatrième chapitre, quant à lui, explore l'utilisation de l'apprentissage profond dans la génération de musique symbolique, et se focalise sur comment réduire les contraintes liées à la tokenisation et au traitement des séquences de grande longueur, tout en améliorant la qualité des compositions générées.

Organisation de la Partie II. La Partie II est organisée comme suit :

1. Chapitre 3 - Évaluation de comportements dynamiques non linéaires : application aux véhicules terrestres.
2. Chapitre 4 - Optimisation des représentations symboliques musicales pour l'apprentissage profond

ÉVALUATION DE COMPORTEMENTS DYNAMIQUES NON LINÉAIRES : APPLICATION AUX VÉHICULES TERRESTRES

On hybrid models

« (...) solving inverse problems with hidden physics is often prohibitively expensive and requires different formulations and elaborate computer codes. Machine learning has emerged as a promising alternative, but training deep neural networks requires big data, not always available for scientific problems. Instead, such networks can be trained from additional information obtained by enforcing the physical laws (...) [Kar+21]

George Em Karniadakis et.al - Nature Reviews, Physics - 2021 »

Sommaire

3.1 Apprentissage informé par la physique : modèles hybrides	110
3.2 Positionnement de nos travaux de recherche	113
3.3 Estimation de paramètres des systèmes dynamiques non linéaires : application à l'évaluation de la dynamique latérale des véhicules	116
3.4 Estimation d'état en temps réel par apprentissage profond et estimation à horizon glissant	136
3.5 Réflexion critique	146

Ce chapitre a pour objectif :

- de rappeler le concept et les principales approches d'apprentissage informé par la physique ;
- de réaliser la synthèse de nos contributions sur l'estimation de paramètres de systèmes dynamiques ;
- de mettre en œuvre une discussion et une réflexion critique sur nos propres contributions à ce champ de recherche.

Contexte des contributions de ce chapitre

Historique. Ce chapitre et les travaux qui en résultent sont liés à la thèse Agence Innovation Défense (AID) de Fabien Lioni intégré à l'INRIA Sophia-Antipolis (06) et à la Direction Générale de l'Armement (DGA) de Montreuil-Juigné (49). Cette thèse a émergé à la suite d'une réponse à appel à projet de l'AID en 2022 que j'ai co-portée et qui a abouti à ce que notre sujet soit retenu. Intitulée « *Évaluation de comportement dynamique par intelligence artificielle : Application à l'analyse de la sécurité du comportement dynamique d'un véhicule* », cette thèse est dirigée par Philippe Martinet (INRIA), et j'ai le plaisir de la co-encadrer depuis 2022 avec Sébastien Aubin (DGA).

Équipe. L'équipe d'encadrement constituée était tripartite : l'INRIA pour la direction de thèse, la DGA et le LERIA pour le co-encadrement.

Environnement et organisation. Les travaux de thèse se sont déroulés pour le doctorant durant 18 mois à Sophia-Antipolis et 18 mois à Angers. Cette configuration m'a donc donné l'occasion de réaliser un suivi à distance et un plus rapproché durant la période où le doctorant était à Angers. Les réunions de suivi se déroulaient au fil de l'eau mais nous avons un point récurrent systématique au moins une fois par semaine.

Mon implication et mes contributions. J'ai été impliqué dans le co-encadrement à hauteur de 40%. Cette mission englobait :

- le suivi régulier du doctorant ;
- le conseil scientifique et stratégique ;
- la participation à l'écriture et la revue d'articles ;
- la participation aux conférences dans lesquelles nous avons publié.

Note importante. Les sections de ce chapitre s'appuient sur sur les travaux suivants : [Lio+24] et [Lio+23] ainsi que

Ces travaux ont été réalisés en collaboration avec les co-auteurs et les contributions intellectuelles sont ainsi partagées entre eux. Ainsi, certaines sections de ce chapitre peuvent provenir de traductions, de reformulations ou encore de formulations textuelles de ces travaux originaux.

3.1 Apprentissage informé par la physique : modèles hybrides

Cette section rappelle un rapide état de l'art général sur les méthodes d'apprentissage informé par la physique et les modèles hybrides. Pour ce faire, elle repose sur des éléments présentés dans [Lio+23 ; Lio+24] ainsi que nos articles acceptés mais en cours de publication. Ils sont accessibles en prépublication ici : <https://ngutowski.fr/article1.pdf> et <https://ngutowski.fr/article2.pdf>.

L'apprentissage informé par la physique (*Physics-Informed Learning*) [Kar+21], et plus précisément les modèles hybrides [LK16], combine des connaissances a priori issues des mo-

dèles physiques avec les capacités d'apprentissage offertes par les approches basées sur les données. Les modèles hybrides combinant des approches basées sur les données et sur les modèles ont montré un grand potentiel pour prédire avec précision les comportements des systèmes, notamment dans des environnements complexes où les modèles traditionnels seuls ne parviennent pas à capturer toutes les dynamiques. Cette synergie entre deux paradigmes permet de développer des modèles prédictifs robustes et interprétables [LK16], particulièrement utiles dans le contexte de systèmes complexes comme par exemple les véhicules autonomes ou les systèmes dynamiques non linéaires.

3.1.1 Modèles d'apprentissage informés par la physique

Dans des environnements complexes, où les phénomènes physiques sont régis par des lois dynamiques non linéaires, les modèles reposant uniquement sur les données (*Model-Based*) peuvent rencontrer des difficultés à généraliser de manière efficace [RPK19]. D'autre part, les modèles physiques peuvent manquer de précision lorsqu'ils font face à des incertitudes ou à des conditions réelles difficiles à modéliser [Wil+20]. C'est à ce titre que les modèles hybrides peuvent prendre tout leur sens en permettant de combler ces lacunes, améliorant ainsi à la fois la précision et la robustesse des prédictions effectuées. Pour ce faire, ces modèles intègrent des contraintes physiques dans des algorithmes d'apprentissage automatique.

L'un des points importants des modèles hybrides réside ainsi en leur capacité à intégrer des mécanismes de robustesse face aux incertitudes et aux bruits dans les données. Par exemple, ceci peut être tout à fait pertinent dans le cadre de l'estimation d'état, où des méthodes comme l'Estimation à Horizon Glissant (*Moving Horizon Estimation (MHE)*) sont employées pour estimer les états d'un système à partir de données bruitées tout en respectant les contraintes physiques du système [RRM03]. En effet, l'entraînement autoregressif avec une fonction de perte intégrale (souvent utilisé dans les modèles reposant sur les données – *data-based*), peut être vu comme une approche identique à celle de la *MHE*, où l'on cherche à minimiser une erreur cumulée sur un horizon temporel glissant. Cela octroie aux modèles hybrides d'obtenir une meilleure robustesse, en leur permettant de mieux gérer les perturbations dans des environnements complexes et dynamiques.

L'une des principales motivations pour utiliser des modèles hybrides [Yin23] repose sur leur capacité à capturer des dynamiques complexes tout en conservant un certain degré d'interprétabilité. En effet, ces modèles exploitent les lois physiques pour guider l'apprentissage et restreindre l'espace de recherche, ce qui réduit le risque de surapprentissage (*overfitting*) tout en rendant les résultats plus explicables [Rud+17].

De plus, il existe d'autres avantages à utiliser des modèles hybrides : 1) En combinant des connaissances a priori et des données expérimentales, les modèles hybrides améliorent la capacité de généralisation dans le cas où les données sont limitées ; 2) L'intégration des lois physiques garantit que les prédictions générées par le modèle respectent les contraintes « naturelles » du système physique, ce qui est primordial dans des applications pouvant être considérées comme critiques p.ex., l'analyse de la dynamique des véhicules.

Une autre problématique majeure pour ces modèles réside dans l'équilibre entre la flexibilité des méthodes basées sur les données (*data-based*) et le respect des contraintes physiques strictes. Des approches comme le *transfer learning*, que nous proposons dans nos travaux, permettent d'introduire ces contraintes de manière "souple", en offrant une certaine flexibilité. Cependant, ces approches peuvent être sensibles aux biais présents dans les données. Il est donc crucial de développer des techniques qui garantissent que les modèles hybrides respectent toujours les lois physiques, même en présence de biais ou de bruit [Zhu+19]. En ce sens, les modèles hybrides doivent être vus comme une forme d'augmentation de données, où les connaissances issues des modèles physiques viennent enrichir les données d'apprentissage pour améliorer la généralisation. Des outils tels que les *multiplicateurs de Lagrange* et les *lagrangiens augmentés* offrent des solutions pour imposer des contraintes rigides dans l'entraînement des modèles [AH24]. Toutefois, ces méthodes ne sont efficaces que si l'on sait précisément où le modèle physique a priori fournit une information plus fiable que les données d'apprentissage. Sinon, l'introduction de contraintes strictes peut, au contraire, dégrader les performances du modèle, surtout lorsque le modèle physique a priori est biaisé.

Enfin, ces modèles réduisent l'incertitude associée aux prédictions, car ils bénéficient à la fois de la robustesse des modèles basés sur des principes physiques et de la flexibilité des méthodes d'apprentissage automatique pour ajuster les paramètres à des données réelles et souvent bruitées. Cela en fait un outil puissant pour la prise de décision en temps réel dans des environnements où la sécurité et la fiabilité sont primordiales, comme dans notre cas les essais de véhicules militaires. De ce fait, les modèles hybrides informés par la physique constituent une approche prometteuse pour la prochaine génération d'outils d'évaluation des systèmes dynamiques.

Les modèles hybrides informés par la physique confèrent donc une approche plus robuste et interprétable. Ceci peut être considéré comme particulièrement pertinent pour des cas d'études sur des systèmes physiques jugés complexes. En intégrant conjointement des méthodes d'apprentissage reposant sur les données et des lois physiques tout en exploitant des techniques inspirées de l'estimation d'état comme la *MHE*, ces modèles peuvent être bien adaptés à des environnements dynamiques. Les travaux en cours dans le cadre de notre collaboration avec notre doctorant visent à approfondir cette exploration. Pour ce faire nous expérimentons différentes méthodes hybrides versus des méthodes reposant uniquement sur les données, tout en évaluant l'impact des contraintes physiques sur la performance des modèles. Nous cherchons à déterminer quelles approches, entre celles intégrant des lois physiques et celles purement basées sur les données, obtiendront les meilleurs résultats.

3.1.2 Applications dans le domaine de la dynamique des véhicules

Dans le cadre de nos travaux sur l'analyse du comportement dynamique des véhicules militaires, les modèles hybrides jouent un rôle crucial pour évaluer leur performance, leur sécurité et leur confort. Par exemple, la modélisation du comportement dynamique d'un véhicule repose sur des équations physiques représentant les interactions entre le véhicule et son environne-

ment (forces d'inertie, friction, etc.) [WYH23]. Cependant, ces modèles peuvent être enrichis par des données provenant de capteurs en temps réel, tels que des centrales inertielles, des caméras ou des accéléromètres [Kue+17].

Grâce à l'intégration de données en temps réel et de modèles physiques pré-établis, l'apprentissage guidé par la physique permet d'améliorer la précision des prédictions tout en assurant que les prédictions respectent les lois physiques fondamentales [Kar+21]. Ces approches ont démontré leur efficacité, surpassant les méthodes purement basées sur les données ou sur les modèles dans des environnements complexes ou non linéaires [Kar+21].

Dans notre cadre applicatif, l'élaboration d'approches hybrides prend tout son sens, mettant ainsi en valeur l'expertise et les connaissances. Il est crucial de reconnaître l'importance de la connaissance a priori sans pour autant négliger celle nécessaire à la découverte. Notre stratégie pour atteindre nos objectifs scientifiques consiste à créer un pont (voir illustration Figure 3.1) entre les approches de l'intelligence artificielle basées sur les modèles et celles basées sur les données. Nous visons à enrichir les modèles avec les connaissances fournies par les approches basées sur les données (*data-based*) (boucle bleue de la Figure 3.1) et, inversement, à contraindre les approches axées sur les données avec des connaissances précises issues de modèles mécanistes et métiers (model-based) (boucle orange de la Figure 3.1).

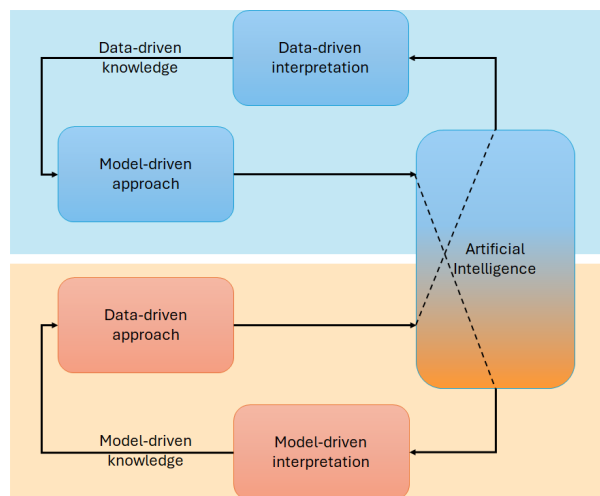


FIGURE 3.1 – Modèle hybride

3.2 Positionnement de nos travaux de recherche

Nos travaux de recherche s'inscrivent dans la convergence de la modélisation des systèmes dynamiques et de l'apprentissage, plus précisément dans le cadre de l'estimation d'états et de paramètres complexes pour des systèmes non-linéaires soumis à des perturbations et des incertitudes. Nous avons formulé trois contributions majeures dans ce contexte, que nous décrivons ci-dessous :

1. Notre première contribution se concentre sur une méthode d'estimation de paramètres basée sur des techniques d'inférence bayésienne pour des systèmes dynamiques non-linéaires. Nous avons développé une approche utilisant l'*Approximate Bayesian Computation (ABC)*, combinée à l'algorithme *Sequential Monte Carlo (SMC)*, pour estimer avec précision les paramètres de modèles complexes décrivant la dynamique de systèmes physiques. Cette méthode intègre des informations a priori sur les distributions paramétriques, permettant ainsi de pallier les limites des méthodes d'optimisation classiques, particulièrement dans les situations où les modèles physiques sont difficiles à calibrer ou où les données sont bruitées et incertaines. L'algorithme *ABC-SMC* se révèle être un outil pertinent pour explorer l'espace des paramètres en simulant différentes configurations, tout en filtrant les solutions qui sont cohérentes avec les observations, même lorsque ces dernières sont incomplètes ou bruitées. Notre approche est particulièrement bien adaptée aux scénarios où les méthodes d'optimisation traditionnelles échouent à capturer la variabilité et l'incertitude inhérente aux systèmes réels. De plus, elle permet de produire des distributions a posteriori sur les paramètres, offrant ainsi une interprétation probabiliste et une meilleure robustesse face aux incertitudes des systèmes non-linéaires. Cette méthode a été appliquée à plusieurs systèmes dynamiques complexes, démontrant des résultats supérieurs en termes de précision et de robustesse des estimations. Nos travaux sur l'estimation des paramètres de la dynamique latérale d'un véhicule ont été publiés à la conférence ISMIS (International Symposium on Methodologies for Intelligent Systems) en 2024 (conférence de rang C).
2. Notre deuxième contribution consiste à proposer une méthode d'apprentissage profond guidée par la physique, combinant des modèles physiques et des techniques d'apprentissage par transfert (Transfer Learning) pour l'estimation des paramètres de la dynamique latérale des véhicules. Cette méthode repose sur une approche hybride qui utilise des modèles physiques bien établis, comme le modèle de bicyclette pour la dynamique latérale et le modèle de Pacejka pour les forces des pneumatiques, pour guider l'entraînement d'un réseau neuronal. Nous résumons son fonctionnement ci-dessous :
 - Dans un premier temps, nous avons utilisé notre approche bayésienne [Lio+24] (sans l'usage de splines), en employant l'algorithme *Approximate Bayesian Computation avec Sequential Monte Carlo (ABC-SMC)*, afin d'estimer les paramètres des modèles physiques. Ces paramètres incluent notamment les coefficients de rigidité des pneus, les distances du centre de gravité, ainsi que d'autres variables influençant le comportement dynamique du véhicule. Cette estimation bayésienne permet d'intégrer des connaissances a priori sur les caractéristiques physiques du véhicule, tout en prenant en compte les incertitudes et les données bruitées.
 - Ensuite, les paramètres estimés sont utilisés pour préentraîner un modèle de réseau neuronal. Ce réseau est ensuite affiné à l'aide de données simulées et réelles dans un cadre d'apprentissage par transfert, ce qui permet d'améliorer la précision de la simulation à long terme dans des conditions dynamiques perturbées, telles que des changements brusques de direction ou des environnements de conduite difficiles.

— Enfin, le réseau préentraîné est appliqué à des données réelles issues de campagnes d'essais sur des véhicules similaires. Cette approche d'apprentissage par transfert assure une généralisation plus robuste du modèle dans des scénarios où les données réelles sont rares ou bruitées, tout en conservant les avantages d'une modélisation physique.

La contribution clé de notre méthode réside dans l'utilisation du préentraînement basé sur des modèles physiques, permettant au réseau neuronal de capturer des comportements dynamiques non observables avec les méthodes traditionnellement basées sur les données. Ces travaux seront publiés dans le journal of Intelligent Information Systems (Springer) (revue de second quartile (Q2) en intelligence artificielle selon le Scimago Journal & Country Rank). Notre article a été accepté le 21/03/2025 mais est en cours de publication à date d'écriture de ce mémoire.

3. Notre troisième contribution repose sur l'utilisation de méthodes d'apprentissage profond pour résoudre des problèmes d'estimation d'état dans des systèmes dynamiques en temps réel. Nous avons développé une méthode de type *Moving Horizon Estimation* (*MHE*, Estimation à horizon glissant) optimisée à l'aide de réseaux de neurones profonds, qui apprend à inférer directement la solution d'un problème d'estimation complexe à partir de données partielles et bruitées. Contrairement à la formulation classique de la *MHE*, notre méthode repose sur une formulation entièrement différentiable de l'algorithme, permettant ainsi d'entraîner le modèle hors ligne en minimisant directement le critère d'optimisation de la *MHE* via des techniques d'apprentissage supervisé. Cette approche permet non seulement de réduire de manière significative les coûts computationnels associés à l'estimation d'état, mais aussi d'améliorer la précision des estimations, même dans des situations avec des modèles imparfaits ou des données bruitées. Une contribution supplémentaire de cette méthode est qu'elle élimine la nécessité d'ajouter des termes de coûts d'arrivée, utilisés dans les formulations classiques de la *MHE* pour garantir la cohérence temporelle des estimations successives. Notre méthode apprend directement à préserver cette cohérence par l'entraînement sur l'ensemble des séquences temporelles. Enfin, cette méthodologie a été validée expérimentalement sur des systèmes dynamiques non linéaires, tels que les oscillateurs de Van der Pol et de Duffing, ainsi que sur des modèles de dynamique latérale de véhicules, avec des résultats montrant une amélioration substantielle par rapport aux méthodes de *MHE* classiques en termes de précision et de temps de calcul. Nos travaux sur l'apprentissage profond pour résoudre des problèmes d'estimation d'état dans des systèmes dynamiques en temps réel seront publiés à la conférence ICRA (IEEE International Conference on Robotics & Automation) en mai 2025 (conférence de rang A*). Notre article a été accepté le 27/01/2025 mais est en cours de publication à date d'écriture de ce mémoire.

Ces trois contributions s'inscrivent dans une démarche visant à améliorer la simulation et l'estimation dans des systèmes dynamiques complexes, en combinant des méthodes d'apprentissage automatique avec des modèles physiques rigoureux. La première contribution permet

de mieux comprendre et estimer les paramètres non observables grâce à des techniques bayésiennes, en utilisant l'*Approximate Bayesian Computation* combiné à l'algorithme *Sequential Monte Carlo* (ABC-SMC). La deuxième contribution se concentre sur l'amélioration de la précision des simulations de dynamiques à long terme en s'appuyant sur des modèles hybrides qui intègrent des connaissances physiques (simulées et générées avec la technique de notre première contribution) avec des méthodes d'apprentissage par transfert. La troisième contribution apporte une solution pour l'estimation d'état en temps réel via des réseaux de neurones profonds, en développant une méthode d'estimation à horizon glissant (MHE) optimisée par apprentissage profond. Ensemble, ces travaux contribuent à réduire l'écart entre les modèles purement physiques et les approches basées sur l'apprentissage automatique, offrant des solutions robustes dans des environnements incertains et partiellement observés.

3.3 Estimation de paramètres des systèmes dynamiques non linéaires : application à l'évaluation de la dynamique latérale des véhicules

Cette section vise à décrire les méthodes d'estimation de paramètres que nous avons mises en place et à rappeler des modèles physiques et les éléments de l'état de l'art de ce domaine. Pour ce faire, elle repose sur des éléments présentés dans [Lio+23; Lio+24] et une extension dans le *Journal of Intelligent Information Systems*, actuellement acceptée, en cours de publication à date d'écriture de ce mémoire (voir la prépublication : <https://ngutowski.fr/article1.pdf>).

3.3.1 Modélisation de la dynamique latérale d'un véhicule

3.3.1.1 Modèle bicyclette à deux degrés de liberté

Le modèle bicyclette à deux degrés de liberté est une représentation simplifiée de la dynamique latérale d'un véhicule. Il est largement utilisé dans le domaine de la dynamique des véhicules pour analyser le comportement de conduite, la stabilité et la maniabilité des véhicules. En appliquant le principe fondamental de la dynamique au modèle des forces décrit Figure 3.2, on obtient le système d'équations différentielles suivant :

$$\dot{v}_y = \frac{1}{m} (F_{yf} \cos(\delta) + F_{yr} - mv_x \dot{\psi}) \quad (3.1)$$

$$\ddot{\psi} = \frac{1}{I_z} (F_{yf} \cos(\delta) - F_{yr}) \quad (3.2)$$

où :

- v_y est la vitesse latérale du véhicule.
- $\dot{\psi}$ est la vitesse de lacet du véhicule.

- F_f et F_r sont les forces latérales aux pneus avant et arrière.
- m est la masse du véhicule.
- v_x est la vitesse longitudinale du véhicule.
- l_f et l_r sont les distances entre le centre de gravité (CG) et les essieux avant et arrière.
- I_z est le moment d'inertie en lacet du véhicule.
- δ est l'angle de braquage des roues avant.

Les forces latérales exercées par les pneus avant et arrière du véhicule sont modélisées à l'aide du modèle non linéaire de Pacejka [Pac+06].

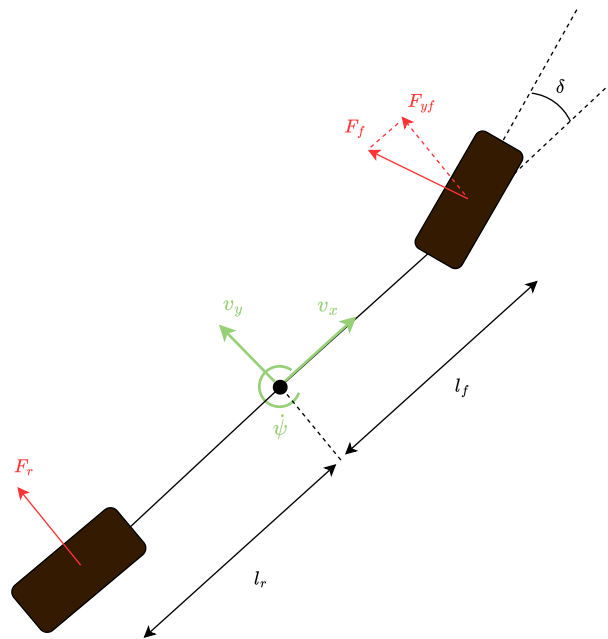


FIGURE 3.2 – Modèle bicyclette à deux degrés de liberté.

Notons enfin que la vitesse latérale v_y et de lacet ψ permettent à eux seuls d'évaluer le comportement dynamique latéral d'un véhicule.

3.3.1.2 Modèle de Pacejka

Le modèle de Pacejka est un modèle dit hybride dans la mesure où il a été établi à partir de lois physiques en supposant une répartition parabolique des forces sur la zone de contact pneu/sol, mais en se fondant également sur des observations phénoménologiques obtenues à partir de résultats d'essais. Ainsi, il s'agit d'un compromis entre un temps de calcul assez rapide et une précision acceptable. Le modèle de Pacejka consiste à modéliser les forces exercées par la route sur les pneus. Les forces sur les pneus F_{y_k} , avec l'indice $k \in \{f, r\}$ correspondant aux roues avant et arrière sont définies comme suit :

$$F_{y_k} = D_k \sin \left(C_k \tan^{-1} (B_k \alpha_k - E_k (B_k \alpha_k - \tan^{-1} (B_k \alpha_k))) \right) \quad (3.3)$$

où :

- B_k est le coefficient de rigidité.
- C_k est le coefficient de forme.
- D_k est la valeur maximale atteinte par F_{y_k} .
- E_k est le facteur de courbure.
- α_k quand $k = f$ est l'angle de dérive du pneu avant et α_k quand $k = r$ l'angle de dérive du pneu arrière.

Le modèle de Pacejka prend en entrée la dérive des pneumatiques avant α_f et arrière α_r , correspondant à la différence entre le vecteur vitesse réel v des pneumatiques et v_y (Figure 3.3). Cet angle de dérive est influencé par des facteurs tels que la vitesse latérale du véhicule, la vitesse de rotation, la distance entre l'essieu avant/arrière et le centre de gravité du véhicule, ainsi que l'angle de braquage des roues avant :

$$\alpha_f = \delta - \tan^{-1} \left(\frac{v_y + l_f \dot{\psi}}{v_x} \right) \quad (3.4)$$

$$\alpha_r = - \tan^{-1} \left(\frac{v_y - l_r \dot{\psi}}{v_x} \right) \quad (3.5)$$

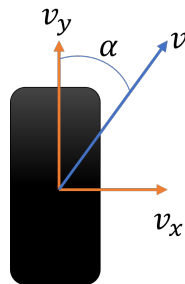


FIGURE 3.3 – Dérive latérale.

3.3.2 Estimation bayésienne des paramètres de systèmes dynamiques non linéaires

Cette sous-section fait référence et décrit notre contribution publiée à la conférence IS-MIS [Lio+24].

3.3.2.1 Estimation des paramètres de systèmes dynamiques non linéaires : vers une approche bayésienne

L'estimation des paramètres de systèmes dynamiques non linéaires est un domaine de recherche déjà largement exploré [Sin89]. L'une des approches classiques pour résoudre ce problème consiste à minimiser une fonction objectif qui mesure l'écart entre les prédictions du modèle et les données observées, en utilisant des méthodes de descente de gradient [BDB86]. Cependant, ce type d'approche présente des limitations comme le fait de converger vers des minima locaux plutôt que le minimum global, ainsi qu'une sensibilité importante aux conditions initiales.

Pour surmonter ces limitations, diverses techniques ont été proposées. L'utilisation d'algorithmes d'optimisation globale tels que les algorithmes génétiques [WX06], le recuit simulé [Eft+02], l'optimisation par essaim de particules [Sch+08] ou *CMA-ES* (*Covariance Matrix Adaptation Evolution Strategy*) [Han16 ; Jeb+07] permet d'explorer plus efficacement l'espace des paramètres à la recherche du minimum global (Voir Chapitre 2). Toutefois, ces approches sont limitées par le nombre de paramètres à estimer : lorsque ce nombre augmente, les temps de calcul augmentent de manière exponentielle. De plus, l'utilisation de méthodes d'intégration numériques, comme la méthode de Runge-Kutta [AP98], pour comparer les prédictions aux observations, entraîne des temps de calcul significatifs, car il est nécessaire d'intégrer le système pour chaque condition initiale étudiée.

Dans le cas d'observations bruitées, il peut être difficile de déterminer avec précision les conditions initiales d'intégration, ce qui peut entraîner des estimations biaisées lorsque le système est sensible à ces conditions. Pour réduire cette sensibilité et éliminer l'étape d'intégration, une approche consiste à lisser les observations à l'aide de splines [AP98], permettant ainsi d'estimer les dérivées des observations du système. Il devient alors possible d'estimer les paramètres à partir du résidu de l'équation différentielle. Ce principe a été récemment étendu aux réseaux de neurones. Les *Physics-Informed Neural Networks (PINN)* [RPK19] permettent d'interpoler des observations correspondant à un processus décrit par des équations aux dérivées partielles (*EDP*), tout en estimant les paramètres de ces équations. Cela est réalisé en s'appuyant sur la différenciation automatique (*Automatic Differentiation - AD*) afin d'optimiser les paramètres par descente de gradient, en minimisant l'erreur quadratique du résidu de l'*EDP*.

Cependant, dans le domaine de l'approximation de paramètres, les approches précédemment mentionnées présentent la limitation de ne pas offrir une interprétation probabiliste des paramètres estimés. Des travaux récents ont cherché à combler cette lacune en incorporant des connaissances sur les distributions *a priori* des paramètres. Ces travaux s'inscrivent dans le cadre bayésien, offrant ainsi deux avantages : restreindre l'espace de recherche des paramètres grâce aux *a priori* spécifiés, et améliorer l'interprétabilité en obtenant une distribution *a posteriori* sur les paramètres. L'approche décrite par [BW14] et [CGL08] consiste à utiliser un processus gaussien pour interpoler et modéliser l'incertitude liée au bruit des observations d'un système dynamique. Étant donné que la dérivée d'un processus gaussien est elle-même un processus gaussien [RW06], il est possible de définir une fonction de vraisemblance en

combinant ce processus avec des méthodes d'échantillonnage, telles que les algorithmes de Metropolis-Hastings [CG95] ou de Gibbs [CG92]. Cette stratégie permet d'estimer une distribution *a posteriori* des paramètres, en supposant une erreur gaussienne entre les dérivées observées et celles prédites par le processus gaussien.

La difficulté de ces méthodes réside dans le paramétrage exigeant des méthodes d'échantillonnage, notamment le choix du nombre d'échantillons pour approximer la distribution *a posteriori*, ou le choix de la proposition de transition. Un mauvais choix de la proposition de transition peut avoir un impact important sur l'efficacité et la performance des méthodes, entraînant un faible taux d'acceptation ou une exploration inefficace de l'espace des paramètres. Il est souvent nécessaire de tester et d'ajuster différentes propositions pour optimiser les performances. De plus, la convergence des méthodes d'échantillonnage est une préoccupation majeure. Il est essentiel que les échantillons générés convergent vers la distribution cible. Cependant, dans certains cas, les méthodes peuvent converger lentement, en particulier lorsque la distribution cible présente des propriétés complexes telles que des modes multiples ou des régions de faible probabilité.

En parallèle, des méthodes *likelihood-free* ont été développées afin d'approximer directement la distribution *a posteriori* sans avoir besoin de définir une fonction de vraisemblance, souvent difficile à établir ou coûteuse à évaluer. Dans ce contexte, la méthode *Approximate Bayesian Computation (ABC)* [Mar+12] permet de fournir une approximation de la distribution *a posteriori* des paramètres en utilisant un échantillonnage basé sur la simulation. Cette approche repose sur un filtrage des paramètres susceptibles d'avoir généré les données observées, en utilisant une mesure de similarité entre les données simulées et les données observées. L'avantage de cette méthode réside dans sa simplicité d'utilisation, car elle ne requiert pas de paramétrages complexes. Pour améliorer l'efficacité de l'*ABC*, l'approche *ABC-SMC* [Ton+09], pour *Sequential Monte Carlo*, augmente l'efficacité de l'échantillonnage en minimisant le nombre de propositions rejetées grâce à un échantillonnage préférentiel qui se concentre sur les zones de l'espace des paramètres les plus probables au regard des données observées.

Notre revue de l'état de l'art n'a pas permis d'identifier d'application de ces méthodes à l'estimation de paramètres dans le cadre de systèmes dynamiques non linéaires de véhicules. Cependant, ces méthodes ont démontré leur efficacité dans l'estimation de paramètres de systèmes dynamiques liés à des processus biologiques [Lie+14 ; LK10 ; STS09]. Dans notre approche, nous évaluons l'utilisation de l'*ABC-SMC* sur des données de simulation s'appuyant sur un modèle de bicyclette intégrant le modèle de pneumatiques de Pacejka [Pac+06]. La problématique étudiée concerne l'estimation des coefficients de Pacejka, ainsi que de certains paramètres physiques du modèle tels que la position du centre de gravité et le moment d'inertie du véhicule. La méthode proposée repose sur le lissage des observations par splines et l'application de l'*ABC-SMC* pour estimer la distribution *a posteriori* des paramètres.

Nos travaux s'ancrent dans une démarche globale et visent à obtenir une modélisation précise de véhicules terrestres militaires. Ceci permet ainsi d'étudier leur stabilité et de garantir un comportement sécuritaire. De plus, l'utilisation d'un modèle numérique fiable permettrait de

réduire les coûts associés aux campagnes d'essais, qui nécessitent actuellement de tester le véhicule dans de nombreux environnements et sous différentes configurations.

De plus, notre contribution est préliminaire à la méthode de transfer learning que nous proposons à la Section 3.3.3 puisqu'elle en devient un composant à part entière.

3.3.2.2 Cadre bayésien

Le cadre bayésien offre une approche idéale pour l'inférence des paramètres des systèmes dynamiques lorsque plusieurs modèles peuvent potentiellement correspondre aux données observées. Dans cette approche, l'ensemble de données \mathcal{D} représente les observations d'un système dynamique, tandis que $\theta = \theta_1, \theta_2, \dots, \theta_n$ désigne l'ensemble des paramètres candidats d'un système d'équations différentielles (ED).

L'équation de Bayes permet de calculer la probabilité *a posteriori* des paramètres θ en fonction des observations \mathcal{D} , comme suit :

$$P(\theta|\mathcal{D}) = \frac{P(\mathcal{D}|\theta) \cdot P(\theta)}{P(\mathcal{D})} \quad (3.6)$$

où :

- $P(\theta|\mathcal{D})$ représente la probabilité *a posteriori* des paramètres θ étant données les données \mathcal{D} .
- $P(\mathcal{D}|\theta)$ est la vraisemblance des données \mathcal{D} étant données les paramètres θ .
- $P(\theta)$ est la probabilité *a priori* des paramètres θ .
- $P(\mathcal{D})$ est la probabilité marginale des données \mathcal{D} .

L'équation de Bayes permet de mettre à jour nos connaissances sur les paramètres du modèle en fonction des données observées et offre un cadre systématique pour incorporer des connaissances *a priori* dans le processus d'inférence.

Le cadre bayésien nécessite d'établir la fonction de vraisemblance $P(\mathcal{D}|\theta)$, ce qui peut poser plusieurs problématiques dans le cas des systèmes dynamiques non linéaires. En effet, ceux-ci sont modélisés par des équations différentielles complexes. Établir une fonction de vraisemblance analytique peut de ce fait être difficile, voire impossible. Cela nécessite donc d'utiliser des techniques d'approximation ou des méthodes numériques pour estimer la vraisemblance. La probabilité marginale $P(\mathcal{D})$ peut être ainsi calculée en intégrant la vraisemblance $P(\mathcal{D}|\theta)$ sur l'ensemble de l'espace des paramètres θ , pondéré par la probabilité *a priori* $P(\theta)$:

$$P(\mathcal{D}) = \int P(\mathcal{D}|\theta) \cdot P(\theta) d\theta \quad (3.7)$$

Dans de nombreux cas, cette intégrale n'a pas de forme analytique connue et ne peut être calculée directement. C'est particulièrement le cas lorsque le modèle est non linéaire ou que la dimension des paramètres est élevée.

3.3.2.3 Approximate Bayesian Computation (ABC)

La méthode ABC [Mar+12] (Voir Algorithme 5) permet d'estimer la distribution *a posteriori* $P(\theta|\mathcal{D})$ sans avoir à spécifier la fonction de vraisemblance et sans avoir à calculer la probabilité marginale des données. Le seul prérequis est de pouvoir simuler des données \mathcal{D}^* à partir d'un modèle correspondant à différents paramétrages de θ .

Algorithme 5 : Algorithme ABC

Data : Données observées \mathcal{D} , taille de l'échantillon N , seuil d'acceptation ϵ

Result : Approximation de la distribution *a posteriori*

```
1 while  $i \leq N$  do
2   Générer un échantillon  $\tilde{\theta}$  à partir de la distribution a priori;
3   Générer des données simulées  $\mathcal{D}^*$  à partir de  $\tilde{\theta}$ ;
4   Calculer la distance  $\rho(\mathcal{D}, \mathcal{D}^*)$ ;
5   if  $\rho < \epsilon$  then
6      $\theta^i \leftarrow \tilde{\theta}$ ;
7      $i = i + 1$ ;
8   end
9 end
```

Cependant, cette approche reste inefficace dans le choix des échantillons θ au cours des itérations, car elle explore l'espace des paramètres sans favoriser les régions les plus susceptibles d'engendrer une acceptation de l'échantillon. L'approche *ABC-SMC* (voir Algorithme 6) permet d'améliorer l'efficacité de l'échantillonnage en s'appuyant sur le principe d'utiliser efficacement les informations antérieures.

Dans l'*ABC-SMC*, les échantillons de l'itération précédente sont utilisés pour générer de nouveaux échantillons. Cela permet de tirer parti des informations déjà acquises sur la distribution *a posteriori* et de concentrer les efforts de calcul sur les régions les plus prometteuses de l'espace des paramètres. Ainsi, en améliorant la qualité de l'échantillonnage, il est possible d'obtenir une amélioration progressive de l'approximation de la distribution *a posteriori* en utilisant une séquence de seuils décroissants de plus en plus stricts.

La génération de nouveaux échantillons repose sur un mécanisme d'échantillonnage par importance des précédents échantillons en pondérant les chances de tirage. Chaque échantillon tiré est alors perturbé par une fonction de perturbation K , souvent uniforme ou gaussienne, qui permet de modéliser la probabilité de transition $P(\theta|\tilde{\theta}_{t-1})$. Elle peut être paramétrée pour contrôler l'amplitude des perturbations et ainsi influencer la diversité et l'exploration de l'espace des paramètres.

Après génération de nouveaux échantillons, ils sont évalués à l'aide de la mesure d'adéquation ρ et les échantillons générant les données \mathcal{D}^* les plus proches des observations \mathcal{D} sont sélectionnés selon le seuil ϵ_t pour la prochaine itération du processus d'échantillonnage. Cette étape itérative est répétée jusqu'à ce que la convergence de la distribution *a posteriori* soit atteinte ou que d'autres critères de convergence soient satisfaits.

Algorithme 6 : Algorithme ABC-SMC

Data : Données observées \mathcal{D} , nombre d'itérations T , taille de l'échantillon N , seuils

$$\epsilon_1 > \epsilon_2 > \dots > \epsilon_T$$

Result : Approximation de la distribution *a posteriori*

```

1 Initialiser l'échantillon  $\theta^1, \theta^2, \dots, \theta^N$ ;
2 for  $t \leftarrow 1$  to  $T$  do
3   while  $i \leq N$  do
4     if  $t = 1$  then
5       |  $\tilde{\theta} \leftarrow P(\theta)$ ;
6     else
7       | Tirer un échantillon  $\tilde{\theta}_{t-1}$  à partir de l'ensemble des échantillons de l'itération
8       | précédente  $\theta_{t-1}$  pondéré par  $w$ ;
9       |  $\tilde{\theta} \leftarrow P(\theta|\tilde{\theta}_{t-1})$ ;
10    end
11    if  $P(\tilde{\theta}) \neq 0$  then
12      | Générer des données simulées  $\mathcal{D}^*$  à partir de  $\tilde{\theta}$ ;
13      | Calculer la distance  $\rho(\mathcal{D}, \mathcal{D}^*)$ ;
14      | if  $\rho < \epsilon_t$  then
15        |  $\theta_t^i \leftarrow \tilde{\theta}$ ;
16        | if  $t \neq 1$  then
17          |  $w_t^i = \frac{P(\theta_t^i)}{\sum_{j=1}^N w_{t-1}^j K(\theta_t^i, \theta_{t-1}^j)}$ ;
18        | else
19          |  $w_t^i = 1$ ;
20        | end
21        |  $i = i + 1$ ;
22      | end
23    end
24  for  $j \leftarrow 1$  to  $N$  do
25    |  $w_t^j = \frac{w_t^j}{\sum_{i=1}^N w_t^i}$ ;
26  end
27 end

```

3.3.2.4 Notre méthode bayésienne pour l'estimation

Génération des données. À partir du modèle décrit précédemment, les données sont utilisées pour l'estimation des paramètres $\theta = \{B, C, D, E, l_f, l_r, I_z\}$. Le système d'équations différentielles est intégré par la méthode de Runge-Kutta [AP98] avec une période d'échantillon-

nage de 0,01 secondes pour une durée de 60 secondes. Un bruit gaussien est ajouté aux trajectoires obtenues, de moyenne nulle et de variance $\sigma_{v_y}^2 = 2,5 \times 10^{-3}$ pour v_y et $\sigma_\psi^2 = 6,25 \times 10^{-6}$ pour ψ . Les paramètres d'intégration utilisés sont les suivants :

Paramètre	B	C	D	E	L_f	L_r	I_z	m	v_x
Valeur	3	2	1	1	1,5	3	1,9	1500	30

Paramétrage. La méthode développée (Figure 3.4) se déroule en deux étapes. Dans un premier temps, les observations obtenues sont lissées pour les variables d'état du système v_y et ψ à l'aide d'un spline cubique. Afin de filtrer le bruit des observations, le spline est régularisé en pénalisant sa courbure de manière à ne pas interpoler le bruit des observations. Deux fonctions $S_{v_y}(x)$ et $S_\psi(x)$ sont obtenues, permettant d'approximer les dérivées des deux variables d'état du système $\dot{S}_{v_y}(x)$ et $\dot{S}_\psi(x)$ malgré le bruit initialement présent dans les observations. $x \in X$ correspond à la position des observations au cours du temps, $X = \{0,01, 0,02, \dots, 60,0\}$.

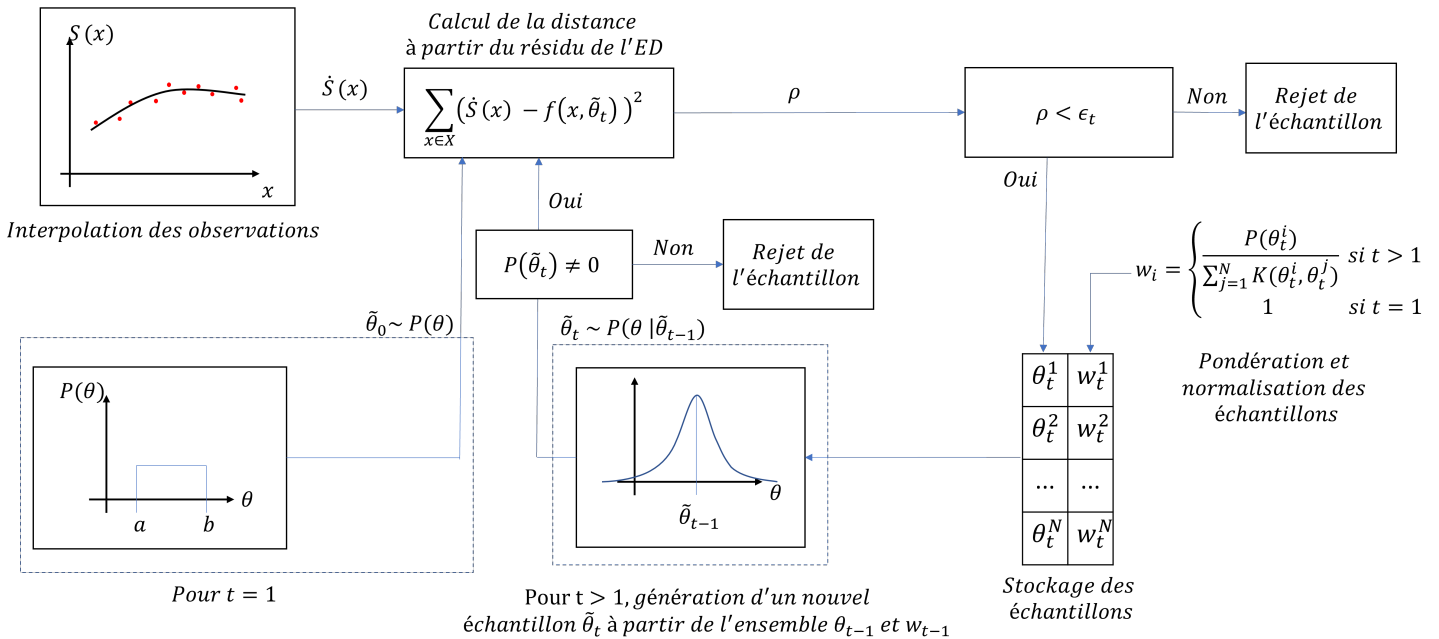


FIGURE 3.4 – Description générale de la méthode : la fonction $S(x)$ correspond à une interpolation réalisée par spline cubique sur les observations d'une variable d'état. Les différentes étapes au cours des itérations t de l'algorithme ABC-SMC y sont représentées. $f(x, \theta_t)$ représente une équation différentielle paramétrée par θ_t .

Le résidu de l'équation différentielle décrivant la dynamique du système est défini de la manière suivante :

$$\begin{cases} \dot{S}_{v_y}(x) - \frac{F_{yf}}{m} \cos \delta_f + \frac{F_{yr}}{m} - v_x S_\psi(x) = 0 \\ \dot{S}_\psi(x) - \frac{l_f}{I_z} F_{yf} \cos \delta_f - \frac{l_r}{I_z} F_{yr} = 0 \end{cases} \quad (3.8)$$

Une approximation de la solution de ce système d'équations différentielles est donnée par :

$$\hat{\theta} = \arg \min_{\theta \in \Theta} \rho(\theta) \quad (3.9)$$

avec :

$$\rho(\theta) = \frac{1}{|X|} \sum_{x \in X} \left(\left(\dot{S}_{v_y}(x) - \frac{F_{yf}}{m} \cos \delta_f + \frac{F_{yr}}{m} - v_x S_\psi(x) \right)^2 + \left(\dot{S}_\psi(x) - \frac{l_f}{I_z} F_{yf} \cos \delta_f - \frac{l_r}{I_z} F_{yr} \right)^2 \right) \quad (3.10)$$

La valeur de $\rho(\theta)$ correspond à une métrique d'adéquation entre les paramètres θ et la solution de l'équation différentielle. $\rho(\theta)$ est utilisée en tant que distance au sein de l'algorithme *ABC-SMC* de manière à rejeter les paramètres qui ont peu de chances de correspondre à une solution du système.

Afin de contraindre la recherche des solutions de l'équation différentielle, un *a priori* uniforme $P(\theta) = \text{Unif}(a, b)$ est spécifié sur chacun des 7 paramètres à estimer :

Paramètre	B	C	D	E	l_f	l_r	I_z
a	2,4	1,4	0,4	0,4	0,9	2,4	1,3
b	4	3	2	2	2,5	4	2,9

Un nombre d'itérations $T = 8$ est spécifié ; pour chacune des itérations, $N = 2000$ échantillons sont générés. Afin de filtrer les échantillons au cours de chacune des itérations, les seuils utilisés sont les suivants : $\epsilon_t = \{0,5; 0,3; 0,15; 0,05; 0,025; 0,02; 0,015; 0,0125\}$.

La fonction de perturbation $P(\theta | \tilde{\theta}_{t-1}) = \mathcal{N}(\mu, \sigma^2)$ utilisée est une gaussienne de variance $\sigma^2 = 0,01$ et de moyenne $\mu = \tilde{\theta}_{t-1}$.

3.3.2.5 Synthèse des résultats

Cette section, synthétise les résultats obtenus dans l'estimation des paramètres de la dynamique latérale d'un véhicule à l'aide de la méthode bayésienne *ABC-SMC* (*Approximate Bayesian Computation - Sequential Monte Carlo*) appliquée à un modèle de bicyclette à deux degrés de liberté, intégrant le modèle de pneumatiques de Pacejka.

Paramètres estimés. Les paramètres estimés incluent ceux du modèle de Pacejka (B, C, D, E) ainsi que les paramètres physiques du véhicule, tels que la longueur des essieux avant

et arrière (l_f, l_r) et le moment d'inertie en lacet (I_z). Les résultats que nous avons obtenus montrent une convergence progressive des échantillons vers la distribution a posteriori. Nous observons plus particulièrement des corrélations importantes entre l_f et l_r , qui sont liés à la position du centre de gravité. Nous observons aussi des corrélations entre les paramètres C et D , correspondant aux caractéristiques des pneus. Les distributions de certains paramètres, comme le facteur de courbure E , présentent une variance plus élevée. Ceci suggère une faible sensibilité du système à ces paramètres dans les conditions expérimentales utilisées.

Analyse des résultats. L'analyse des résultats met en exergue que notre méthode est capable de fournir des estimations fiables et robustes des différents paramètres, et ce même en présence de données bruitées. Les estimations des paramètres B (rigidité des pneus), D (force maximale exercée par les pneus), ainsi que les variables dynamiques $\dot{\psi}$ (vitesse de lacet) et v_y (vitesse latérale), sont cohérentes avec les trajectoires simulées du véhicule, ce qui confirme la capacité de notre méthode à capturer avec fidélité la dynamique latérale du véhicule.

Les figures (résultats) disponibles dans notre article [Lio+24] illustrent la convergence des échantillons au cours des itérations de l'algorithme *ABC-SMC* et montrent une bonne correspondance entre les trajectoires obtenues par notre méthode et les observations initiales bruitées. En effet, les erreurs quadratiques moyennes (RMSE) entre les trajectoires obtenues par l'intégration du système avec les paramètres estimés et les données d'observation montrent des valeurs très proches de celles des données d'origine. Ceci indique que notre méthode parvient à identifier les paramètres efficacement malgré le bruit.

Ainsi, la méthode que nous avons développée dans cette étude a démontré son efficacité pour estimer les paramètres dynamiques d'un véhicule dans un cadre bayésien. Elle est même particulièrement efficace lorsqu'elle est appliquée à des données bruitées et limitées. L'approche *ABC-SMC* a permis d'obtenir une approximation précise de la distribution *a posteriori* des paramètres, fournissant une vision probabiliste qui permet de mieux comprendre la sensibilité et l'incertitude des paramètres estimés. Ces résultats ouvrent des perspectives intéressantes pour des applications dans la modélisation de véhicules ou d'autres systèmes dynamiques complexes. Nous en avons réalisé une application spécifique et nous avons proposé une nouvelle contribution dans ce domaine. Nous les présentons dans la section 3.3.3.

3.3.2.6 Discussion

L'analyse réflexive complémentaire de cette étude sur l'estimation des paramètres dynamiques d'un véhicule par méthode bayésienne *ABC-SMC* permet de mettre en lumière plusieurs points pertinents sur les méthodes utilisées et leur applicabilité dans d'autres contextes.

Pertinence et robustesse de la méthode. Avoir combiné du lissage par spline et de l'*ABC-SMC* présente un grand intérêt dans les systèmes où les données sont limitées ou bruitées. L'approche que nous avons choisie permet de contourner certaines limitations des méthodes

traditionnelles d'estimation des paramètres dynamiques, comme l'intégration numérique répétitive ou les méthodes d'optimisation globale. Celles-ci peuvent notamment devenir très coûteuses en temps de calcul lorsque le nombre de paramètres à estimer augmente. De plus, l'intégration de connaissances a priori nous a permis de restreindre l'espace de recherche, rendant le processus plus efficient et précis. Ceci est important à considérer dans le sens où des essais physiques sur les véhicules restent coûteux.

Cependant, notre méthode présente aussi certaines limites, notamment en ce qui concerne la dépendance à la qualité des a priori et des données d'observation. Si les a priori sont mal spécifiés, cela pourrait biaiser les résultats finaux, limitant la fiabilité de l'estimation. Dans les systèmes fortement non linéaires, les biais potentiels peuvent être amplifiés, surtout si les observations sont elles-mêmes limitées ou si elles contiennent des biais importants.

Interprétabilité et applications pratiques. L'un des points intéressants de notre étude est l'interprétabilité que nous permet l'emploi d'une méthode bayésienne. Contrairement aux approches traditionnelles, qui fournissent une estimation ponctuelle des paramètres, l'approche bayésienne génère une distribution a posteriori, permettant ainsi une évaluation probabiliste des paramètres. Cette évaluation probabiliste permet ainsi de mieux comprendre la sensibilité des différents paramètres du système et de mieux guider la prise de décision dans des contextes où il est primordial de mieux appréhender l'incertitude.

Dans une application pratique, comme l'estimation des paramètres dynamiques d'un véhicule, cette méthode pourrait être utilisée pour des essais de modélisation prédictive. Par exemple, dans les domaines des transports ou comme dans notre cas de la défense, où la modélisation précise de la dynamique est critique pour des tests en conditions extrêmes, cette méthode pourrait offrir un cadre rigoureux pour simuler des scénarios rares ou coûteux, tout en prenant en compte les incertitudes inhérentes aux systèmes réels.

Limites et voies d'amélioration. Bien que l'approche soit efficace, plusieurs pistes d'amélioration peuvent être envisagées :

- L'une des limites identifiées est la faible quantité de données disponibles, ce qui affecte la précision de l'estimation de certains paramètres, comme le facteur de courbure des pneus. En situation réelle, des campagnes d'essais additionnelles pourraient être nécessaires pour affiner ces estimations et réduire la variance observée dans les distributions a posteriori.
- L'approche que nous avons créée repose sur un modèle simplifié de véhicule, comme le modèle bicyclette à deux degrés de liberté. Si cette approche est largement utilisée, elle reste une approximation simplifiée. Il pourrait être intéressant d'étendre cette méthode à des modèles plus complexes, intégrant par exemple la dynamique verticale du véhicule, afin d'évaluer la robustesse de l'approche dans des systèmes plus complexes.
- Bien que l'approche bayésienne évite l'intégration répétée des équations différentielles, elle introduit d'autres sources de complexité, notamment la nécessité d'effectuer un

grand nombre d'itérations et d'échantillonnages pour converger vers une solution acceptable. L'optimisation des algorithmes d'échantillonnage, notamment par l'amélioration des propositions de transition dans *ABC-SMC*, pourrait améliorer encore davantage la performance de la méthode.

3.3.3 Estimation des paramètres de la dynamique des véhicules par apprentissage par transfert guidé par la physique

Cette sous-section fait référence et décrit notre contribution (acceptée) qui sera publiée dans le journal of Intelligent Information Systems (voir la prépublication : <https://ngutowski.fr/article1.pdf>).

3.3.3.1 Estimation des paramètres de la dynamique des véhicules : vers une approche hybride

Les modèles dynamiques jouent un rôle crucial dans la réalisation de simulations numériques, outils essentiels pour diverses tâches d'ingénierie. Ces simulations permettent de prédire et d'analyser le comportement de systèmes complexes dans des conditions variées, sans avoir à recourir à des expériences coûteuses et chronophages en conditions réelles. Cela est indispensable pour la conception et l'optimisation des systèmes, l'évaluation de la sécurité, ainsi que le développement de stratégies de contrôle [Hu+21]. Cependant, obtenir des modèles dynamiques précis, capables de fournir des simulations fiables sur le long terme, reste un enjeu majeur [Jim+17]. La complexité des systèmes réels, combinée aux incertitudes des paramètres du modèle et aux perturbations externes, entraîne souvent des écarts significatifs entre les comportements simulés et observés, particulièrement lorsque les simulations couvrent de longues périodes, en raison d'une faible capacité de généralisation et de la propagation croissante des erreurs.

Méthodes existantes et genèse de notre méthode hybride en trois étapes. L'identification des systèmes est un problème récurrent dans l'ingénierie des systèmes, en particulier pour les systèmes dynamiques non linéaires. Cela implique souvent la résolution de problèmes inverses mal posés, en raison de la multiplicité des ensembles de paramètres possibles correspondant à un modèle physique donné [Lju99]. La plupart des approches classiques reposent sur l'optimisation de problèmes de moindres carrés [Sto98], qui consiste à trouver le jeu de paramètres maximisant la probabilité d'aligner les trajectoires des variables d'état prédites avec les observations, sous une hypothèse d'erreur gaussienne. Cependant, avec des données peu informatives, cette multiplicité peut entraîner des estimations de paramètres restreintes à la plage des mesures disponibles, limitant ainsi la capacité du modèle à généraliser.

Pour résoudre ce problème, les méthodes d'inférence bayésienne [IBK06] sont largement utilisées (voir sous-section 3.3.2.1). Elles permettent d'incorporer des connaissances a priori

sur les distributions des paramètres, agissant ainsi comme des contraintes pour obtenir des estimations physiquement significatives, même avec des données limitées. Parmi ces méthodes, l'*Approximate Bayesian Computation (ABC)* [Mar+12] s'est montrée particulièrement utile pour estimer les probabilités a posteriori des paramètres lorsque la fonction de vraisemblance est intractable, notamment dans des systèmes dynamiques non linéaires et en présence de bruits non gaussiens. Ces méthodes ont été appliquées avec succès dans plusieurs domaines, allant de la modélisation de la dynamique latérale des véhicules [Lio+24] à la calibration de simulations en cosmologie [Ake+15]. **Ces méthodes bayésiennes nous intéressent plus particulièrement, car l'une d'entre elles, que nous avons développée [Lio+24] dans la sous-section 3.3.2.1, est le composant de la première étape et de la seconde étape (voir Figure 3.5) de la méthode hybride que nous proposons dans cette sous-section. Elle permet de pouvoir générer des données sources d'apprentissages simulant et respectant les principes de la physique.**

D'autre part, les approches basées sur l'apprentissage automatique offrent une flexibilité accrue pour modéliser directement des phénomènes dynamiques complexes avec peu d'hypothèses sur le processus sous-jacent. Les perceptrons multicouches (*MLP*) sont connus pour leur capacité d'approximation universelle [Pin99] et ont été utilisés pour l'identification de dynamiques de systèmes [GR98]. Cependant, cette flexibilité exige une quantité importante de données couvrant l'ensemble des espaces d'état et de contrôle pour garantir une bonne capacité de généralisation. Récemment, les architectures de type ResNet [He+16] ont montré des similitudes avec les méthodes d'intégration numérique d'Euler explicite, ce qui permet d'optimiser les architectures de réseaux neuronaux sur plusieurs étapes d'intégration temporelle en utilisant des frameworks de différenciation automatique [Wei17] (*AD - Automatic Differentiation*). Pour améliorer les performances de généralisation des méthodes d'apprentissage, l'apprentissage automatique guidé par la physique cherche à combiner des connaissances physiques a priori avec des modèles basés sur les données. Une approche courante consiste à apprendre directement les écarts entre des modèles physiques traditionnels et les mesures réelles [BO14], permettant ainsi d'intégrer les phénomènes physiques non modélisés (c.-à-d., intégrer la physique manquante) et d'améliorer la précision. D'autres méthodes s'appuient sur des fonctions de perte informées par la physique, qui contraignent le processus d'apprentissage à respecter des principes physiques, tels que la conservation de l'énergie, comme démontré dans les réseaux neuronaux hamiltoniens [GDY19]. En outre, certaines approches intègrent directement des équations physiques dans les architectures neuronales, comme l'équation de Lagrange, ce qui a montré des performances supérieures par rapport aux architectures classiques [Cra+20]. **Ces approches basées sur l'apprentissage automatique et l'apprentissage profond, nous intéressent plus particulièrement car l'une d'entre elles, le *MLP*, constitue le composant de la seconde étape (voir Figure 3.5) de la méthode hybride que nous proposons dans cette sous-section. Nous pré-entraînons ainsi ce *MLP* sur les données sources générées par notre méthode d'inférence bayésienne.**

Parallèlement, l'apprentissage par transfert [WKW16] s'est révélé efficace pour exploiter des connaissances acquises d'une tâche à partir d'un ensemble de données source, afin de

les réutiliser dans une tâche cible, particulièrement lorsque cette dernière dispose de moins de données. Cette approche a été largement explorée dans le domaine de la robotique, notamment pour adapter des domaines de simulation à des applications réelles, avec des améliorations notables en termes de vitesse d'apprentissage dans des contextes robotiques complexes [Bou+18]. **Cette dernière approche nous intéresse puisqu'elle constitue le cœur de notre contribution. Une fois le MLP pré-entraînée, nous l'affinons en utilisant une formulation autorégressive durant la fin de l'étape 3 (voir Figure 3.5), par les données de mesure réelle collectée sur un camion *Renault Premium* de la DGA.**

Vers une approche hybride pour l'estimation des paramètres. L'intégration des méthodes bayésiennes et des approches d'apprentissage automatique guidé par la physique représente une voie prometteuse pour répondre aux problématiques de l'estimation des paramètres dans les systèmes dynamiques complexes. En combinant les forces des modèles basés sur la physique — qui offrent des interprétations claires et nécessitent peu de données — avec les capacités de modélisation flexible des approches basées sur l'apprentissage, il est possible de surmonter les limitations de généralisation observées sur des simulations à long terme. Dans cette optique, l'apprentissage par transfert apparaît comme un cadre adapté pour tirer parti des simulations issues de modèles physiques connus afin d'améliorer les performances sur des données réelles plus limitées.

Notre méthode en trois étapes repose précisément sur ce principe : Nous combinons l'apprentissage par transfert avec des techniques d'inférence bayésienne pour l'estimation des paramètres, en nous appuyant sur des simulations basées sur la physique et sur des mesures du monde réel. Notre approche permet de garantir une meilleure robustesse des modèles tout en minimisant la propagation des erreurs sur le long terme. Les détails de cette méthodologie sont exposés dans la sous-section 3.3.3.2 suivante.

3.3.3.2 Notre méthode hybride en trois étapes

Vue d'ensemble. Dans le cadre de l'apprentissage par transfert, la méthode que nous proposons exploite un ensemble de données source \mathcal{D}_{src} et un ensemble de données cible \mathcal{D} . L'approche consiste à transférer les connaissances provenant des données de simulation \mathcal{D}_{src} , générées à l'aide d'un modèle dynamique non linéaire de bicyclette (voir sous-section 3.3.1.1), vers une tâche cible consistant à apprendre la dynamique latérale du véhicule à partir de données réelles \mathcal{D} . Cela est réalisé par le biais d'une formulation d'entraînement composée de trois étapes principales (voir Figure 3.5) :

- **La première étape** consiste à estimer les paramètres θ à partir du modèle de bicyclette $f(x_t, u_t, \theta)$, où $x_t \in \mathbb{R}^n$ désigne l'état et $u_t \in \mathbb{R}^m$ désigne la commande appliquée à l'instant t . Bien que θ représente un ensemble de paramètres fixes à estimer, des croyances a priori sur leur distribution sont disponibles. Cette estimation des paramètres est réalisée en utilisant l'inférence bayésienne, en s'appuyant sur les données de mesure du système \mathcal{D} et sur le modèle physique a priori de la dynamique latérale f ainsi que sur

les croyances concernant l'intervalle des paramètres $P(\theta)$. En pratique, pour effectuer cette estimation bayésienne, nous utilisons une méthode d'échantillonnage : *ABC-SMC* (comme réalisé en Section 3.3.2.4). L'avantage de *ABC-SMC* en tant qu'approche itérative est qu'elle permet de raffiner progressivement l'estimation des paramètres. Pour ce faire elle échantillonne des ensembles de paramètres et compare les prédictions du modèle avec les données réelles, tout en considérant les incertitudes associées aux paramètres et aux mesures.

- **La deuxième étape** consiste à générer des données de simulation \mathcal{D}_{src} en utilisant f avec l'ensemble de paramètres qui maximise l'a posteriori $P(\theta|\mathcal{D})$.
- Durant **la troisième étape**, cet ensemble de données \mathcal{D}_{src} est utilisé pour pré-entraîner un perceptron multicouche (*MLP*) qui est ensuite affiné sur les données de mesure \mathcal{D} en utilisant une formulation d'entraînement autorégressive.

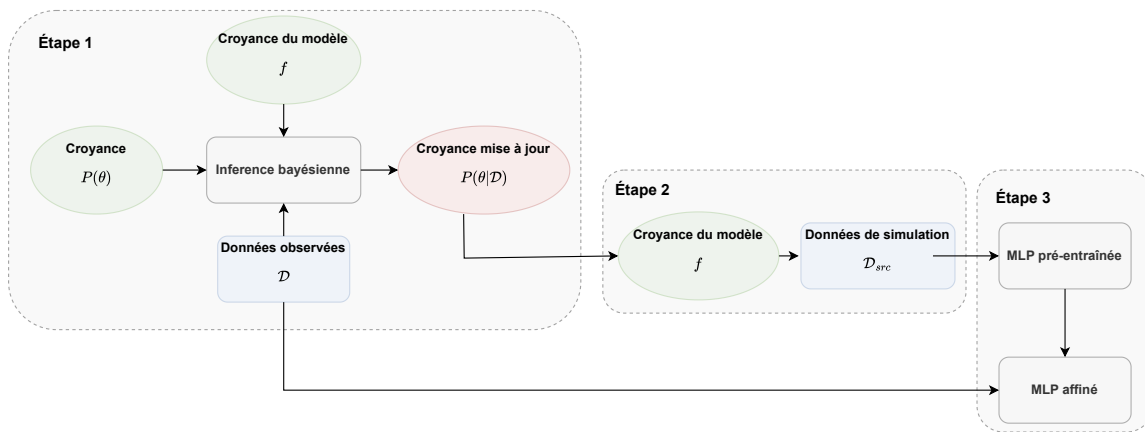


FIGURE 3.5 – Récapitulatif de la méthode proposée

Détails de l'étape 1 sur la partie identification des paramètres avec *ABC-SMC*. La première étape de la méthode proposée consiste à identifier les paramètres θ du modèle de bicyclette en utilisant l'inférence bayésienne. Cette approche permet de combiner des informations a priori $P(\theta)$ sur les paramètres avec une fonction de vraisemblance $P(\mathcal{D}|\theta)$, qui représente la probabilité d'observer l'ensemble de données \mathcal{D} étant donné les paramètres θ . Cette combinaison donne la distribution a posteriori $P(\theta|\mathcal{D})$, qui reflète la croyance mise à jour sur les paramètres après prise en compte des données (Voir Equation 3.6).

Pour réaliser cette tâche d'inférence, nous utilisons de nouveau la variante de la méthode d'*Approximate Bayesian Computation* (*ABC*) appelée *ABC-SMC* (*Sequential Monte Carlo*) [Ton+09] (Voir Algorithme 6), mais cette fois sans lissage des observations par splines.

ABC-SMC repose sur un raffinement itératif d'une population de particules représentant les paramètres du système. La population initiale est échantillonnée à partir de la distribution a priori des paramètres et filtrée de manière itérative en fonction d'un critère de distance ρ .

À chaque itération p , un nouvel ensemble de particules est obtenu par rééchantillonnage des particules acceptées de l'itération précédente, en utilisant un schéma d'échantillonnage d'importance. Les nouvelles particules sont ensuite obtenues en perturbant aléatoirement les particules rééchantillonnées à l'aide d'une distribution basée sur un noyau. En suivant ce principe, les particules convergent progressivement vers une approximation de la distribution a posteriori. Le rôle de la distance ρ est de sélectionner itérativement les paramètres susceptibles de correspondre aux paramètres de la distribution a posteriori. Pour une petite valeur de ϵ , les ensembles de paramètres génèrent des données simulées \mathcal{D}_θ avec le modèle f proches de \mathcal{D} , approximant ainsi la distribution a posteriori :

$$P(\theta|\mathcal{D}) = \lim_{\epsilon \rightarrow 0} P(\rho(\mathcal{D}, \mathcal{D}_\theta) < \epsilon | \theta) \quad (3.11)$$

où ϵ représente le seuil final dans une séquence de P seuils décroissants, notés $\{\epsilon_0, \epsilon_1, \dots, \epsilon_{P-1}, \epsilon\}$. À chaque itération, ces seuils sont appliqués pour guider la sélection des ensembles de paramètres. Ici, \mathcal{D}_θ fait référence à l'ensemble des trajectoires prédites $\{\hat{x}_0^n, \hat{x}_1^n, \dots, \hat{x}_T^n\}$, tandis que \mathcal{D} représente l'ensemble correspondant des trajectoires mesurées des variables d'état $\{x_0^n, x_1^n, \dots, x_T^n\}$. L'indice n désigne différentes trajectoires des variables d'état, chacune couvrant T pas de temps (étapes temporelles) générées à l'aide du schéma d'intégration d'Euler explicite. La constante Δ_t est la taille du pas de temps utilisé dans l'intégration :

$$\begin{cases} \hat{x}_{t+1}^n = \hat{x}_t^n + f(\hat{x}_t^n, u_t^n, \theta)\Delta_t \\ \hat{x}_0^n = x_0^n \end{cases} \quad (3.12)$$

La métrique de distance ρ est définie comme l'erreur quadratique moyenne (RMSE) entre les prédictions et les observations sur T étapes de trajectoires, moyennée sur toutes les N trajectoires :

$$\rho(\mathcal{D}, \mathcal{D}_\theta) = \frac{1}{N} \sqrt{\frac{1}{T} \sum_{t=0}^T \left(\frac{\hat{x}_t^n - x_t^n}{\sigma_x} \right)^2} \quad (3.14)$$

Pour éviter de privilégier une variable d'état en particulier, la distance est standardisée par σ_x , l'écart-type calculé sur chaque variable d'état de \mathcal{D} .

Détail de l'étape 2 sur la génération des données de simulation. Pour générer des données de simulation pour la phase de pré-entraînement, quinze particules correspondant à la dernière population P de l'ABC-SMC qui minimisent ρ sont retenues. Ces particules correspondent au sous-ensemble de paramètres θ_{MAP} qui maximise la distribution a posteriori $P(\theta|\mathcal{D})$. À partir de ces particules sélectionnées, un ensemble de données de simulation \mathcal{D}_{src} est généré. Cela se fait en échantillonnant aléatoirement des variables d'état initiales x_t et des variables de commande u_t à partir de \mathcal{D} , puis en générant des prédictions à une étape x_{t+1} en utilisant la fonction f avec un paramètre θ échantillonné aléatoirement à partir de l'ensemble

θ_{MAP} . L'ensemble de données \mathcal{D}_{src} capture ainsi la variabilité des prédictions de f correspondant à l'incertitude dans l'estimation a posteriori maximale.

Détail de l'étape 3 sur le pré-entraînement du MLP et l'apprentissage par transfert. Nous définissons ensuite l'architecture du perceptron multicouche (*MLP*) de la manière suivante :

$$\dot{x}_t = \phi(x_t, u_t, \beta) \quad (3.15)$$

où β représente les paramètres du modèle *MLP*.

Pendant les phases de pré-entraînement et d'affinage de l'architecture *MLP*, l'optimisation est basée sur la fonction de perte suivante :

$$\mathcal{L}(\beta) = \frac{1}{T} \sum_{t=0}^T (\hat{x}_t - x_t)^2 \quad (3.16)$$

Lors de la phase de pré-entraînement, comme \mathcal{D}_{src} ne contient pas de perturbations, le *MLP* est optimisé pour un seul pas de temps à la fois en fixant $T = 1$. Le *MLP* ne se concentre ainsi que sur la relation immédiate entre les variables d'état et les commandes à un instant donné. Il n'a pas besoin de gérer les erreurs accumulées au fil du temps, car chaque prédiction est faite pour une seule étape. Une fois que le modèle a acquis une bonne capacité à prédire pour un seul pas, il peut ensuite être affiné (fine-tuning) pour gérer des prédictions plus complexes sur plusieurs pas (p.ex., $T=200$ dans notre expérience). Cela permet de progressivement introduire la difficulté d'accumuler des erreurs sur plusieurs étapes. Lors de cette phase d'affinage, le *MLP* commence donc avec la paramétrisation obtenue lors de la phase de pré-entraînement. Pendant cette phase, le processus d'entraînement consiste ainsi à optimiser la génération de trajectoires autorégressives en utilisant ϕ basée sur la méthode d'intégration d'Euler explicite sur de multiples pas de temps T . Ce processus exploite la différenciation automatique des opérations d'intégration numérique impliquées, comme décrit dans [Wei17].

3.3.3.3 Synthèse des résultats

Le détail des résultats est disponible dans notre article soumis à un journal et en cours de relecture à date d'écriture de ce mémoire. Il est accessible au travers de ce lien <https://ngutowski.fr/article1.pdf>.

Les expérimentations menées pour évaluer la méthode d'apprentissage par transfert proposée ont été réalisées en utilisant un ensemble de données collecté sur un camion *Renault Premium* de la DGA. Ce jeu de données \mathcal{D} comprend des mesures de la vitesse longitudinale v_x , la vitesse latérale v_y , et la vitesse de lacet $\dot{\psi}$, obtenues à l'aide de capteurs optiques et d'une unité de mesure inertielle (*IMU - Inertial Measurement Unit*). Les essais ont été effectués sur une piste d'essai avec des vitesses variant de 15 à 80 km/h, et incluaient différentes manœuvres telles que le slalom, les manœuvres d'évitement et les virages à rayon constant. Le jeu de données total couvre environ 20 minutes d'acquisitions.

Notre approche d'apprentissage par transfert a été comparée à plusieurs méthodes, notamment :

1. Le modèle de bicyclette standard sans apprentissage supplémentaire.
2. Un perceptron multicouche (*MLP*) entraîné uniquement sur les données mesurées.
3. Un *MLP* chargé d'apprendre les écarts du modèle de bicyclette, comme décrit dans des travaux antérieurs de [CX21] (c.-à-d., discrepancy, correction des écarts).

Afin de pouvoir bénéficier d'une évaluation équilibrée, nous avons divisé le jeu de données en trois parties : 50% pour l'entraînement, 30% pour la validation, et 20% pour les tests. Nous avons utilisé plusieurs paramètres pour évaluer les performances des méthodes, y compris la complexité des *MLP* définie par la profondeur et le nombre de neurones par couche, ainsi que le temps d'intégration de la fonction de perte lors de la phase d'entraînement.

Pour effectuer l'évaluation nous avons calculé l'erreur quadratique moyenne (RMSE) entre les trajectoires mesurées et prédites sur un horizon de 10 secondes, utilisant un pas de temps de 0.05 seconde, soit $T = 200$ pas d'intégration d'Euler. Les prédictions sont produites à partir d'un état initial x_0 correspondant aux données mesurées, en utilisant les mêmes commandes appliquées que dans les mesures du jeu de test. La RMSE a été utilisée pour quantifier la différence moyenne entre les prédictions et les données réelles.

Les résultats que nous avons obtenus ont montré que la méthode d'apprentissage par transfert surpasse systématiquement les autres méthodes, avec des RMSE mesurées plus faibles et une variabilité réduite pour les états de vitesse latérale v_y et de lacet $\dot{\psi}$. Augmenter la complexité du *MLP*, notamment avec 16 et 24 neurones par couche, a également permis d'améliorer les performances, bien que les gains de précision deviennent moins significatifs au-delà de 16 neurones.

L'estimation des paramètres du modèle de bicyclette par la méthode *ABC-SMC* a été efficace et robuste, comme le montre la réduction de la RMSE avec des temps d'intégration plus longs. Cependant, l'obtention de cette bonne précision a nécessité une puissance de calcul plus importante. La phase de d'apprentissage par transfert a exploité cette estimation pour pré-entraîner ensuite le *MLP*, conduisant à une meilleure généralisation dans les simulations sur le long terme.

De plus, nous avons constaté que les méthodes basées sur l'apprentissage des écarts étaient moins stables en raison de la dépendance des prédictions auto-régressives sur le modèle de bicyclette, entraînant ainsi des gradients instables. Notre méthode d'apprentissage par transfert quant à elle a montré des améliorations notables, notamment lors de manœuvres plutôt complexes comme des slaloms, où la dynamique du véhicule subit des phénomènes de transfert de masse.

En résumé, on observe que la méthode d'apprentissage par transfert proposée a obtenu une meilleure capacité à généraliser sur des horizons temporels longs, en exploitant à la fois des simulations basées sur des modèles physiques et des données réelles.

3.3.3.4 Discussion

Nos travaux se sont situés à l'intersection de l'apprentissage automatique, des méthodes bayésiennes et de la modélisation physique des systèmes dynamiques. Ils mettent en lumière un cadre novateur d'apprentissage par transfert, qui combine la précision des modèles physiques avec la flexibilité des approches d'apprentissage automatique. Cette combinaison permet d'améliorer significativement les capacités de généralisation des modèles sur des horizons temporels longs, tout en prenant en compte les incertitudes et les perturbations inhérentes aux systèmes physiques complexes.

L'un des principaux avantages de nos contributions est la capacité à exploiter à la fois les connaissances a priori sur la dynamique physique des véhicules et les données empiriques. Employer l'inférence bayésienne au travers de l'algorithme *ABC-SMC* pour estimer les paramètres du modèle de bicyclette en s'appuyant sur des distributions a priori et des données mesurées a permis une intégration efficace des modèles théoriques et des observations réelles. Cette approche a permis d'ancrer le processus de modélisation dans une base physique solide, tout en considérant les incertitudes inhérentes aux paramètres.

De plus, l'utilisation de l'apprentissage par transfert a permis au réseau de neurones de pré-apprendre à partir des simulations générées par le modèle physique. Cette approche réduit la dépendance directe aux données réelles, qui peuvent être coûteuses à obtenir ou porteuses d'erreurs. Cette stratégie de pré-entraînement a amélioré la robustesse du modèle et a favorisé la capacité de généraliser, notamment dans des cas où les données de test ont pu différer légèrement de celles observées durant la phase d'entraînement.

Les résultats expérimentaux que nous avons obtenus, ont démontré que notre méthode surpassait les approches classiques comme le modèle de bicyclette seul, les *MLP* sans apprentissage par transfert ou encore les modèles discrepancy (correction des écarts). Cette amélioration s'explique par la capacité de notre méthode à capturer les dynamiques complexes du système tout en limitant la propagation des erreurs dans les prédictions à long terme. L'approche autorégressive utilisée dans la phase de fine-tuning (affinage) a renforcé cette capacité, car elle exploite les sorties du modèle à chaque pas de temps pour générer les prédictions futures.

Limites et réflexion éthique. D'un point de vue éthique, l'application de telles méthodes dans des contextes critiques, comme les véhicules autonomes ou la dynamique des transports, soulève des questions de responsabilité. En effet, le recours à des modèles basés sur l'apprentissage implique que les décisions prises par ces systèmes peuvent ne pas être entièrement explicables, notamment si le réseau de neurones se base sur des dynamiques apprises implicitement. Il est donc essentiel de garantir une transparence dans les systèmes de modélisation pour assurer une compréhension claire des prédictions et des décisions prises par les algorithmes.

Perspectives et généralisabilité. En perspective, nos travaux pourraient être étendus à d'autres types de systèmes dynamiques complexes, au-delà des véhicules terrestres. La combinaison de modèles physiques et d'apprentissage par transfert pourrait trouver des applications dans des domaines variés, tels que la robotique, l'aéronautique ou encore les systèmes biologiques, où la modélisation précise des dynamiques est cruciale et les données expérimentales sont coûteuses ou difficiles à obtenir. De plus, des améliorations pourraient être apportées en explorant des architectures de réseaux de neurones plus complexes ou en intégrant des mécanismes de régularisation basés sur des lois physiques plus générales.

En conclusion, nous pouvons considérer que notre contribution a montré comment les techniques d'apprentissage automatique et de modélisation physique peuvent être harmonieusement combinées pour améliorer la précision et la robustesse des modèles dynamiques complexes, ouvrant ainsi de nouvelles perspectives pour des applications industrielles et scientifiques.

3.4 Estimation d'état en temps réel par apprentissage profond et estimation à horizon glissant

Cette section vise à décrire les méthodes d'estimation d'état que nous avons mises en place et à rappeler des éléments de l'état de l'art de ce domaine. Pour ce faire, elle repose sur un article accepté à la conférence IEEE International Conference on Robotics and Automation (ICRA) le 27/01/2025 (Prépublication : <https://ngutowski.fr/article2.pdf>). Cet article sera disponible dans les proceedings d'ICRA en mai 2025.

3.4.1 Estimation d'état des systèmes dynamiques

L'estimation de l'état des systèmes dynamiques à partir de mesures a toujours été une problématique scientifique majeure en théorie du contrôle. Cette complexité découle de divers facteurs tels que : des contraintes d'intégration dues aux capteurs, des limitations dans des conditions locales extrêmes, l'absence de capteurs pour des mesures spécifiques, ou tout simplement des contraintes de coût. Surmonter ces obstacles est primordial pour le développement de contrôleurs efficaces, en particulier lorsque l'on traite d'états qui ne peuvent être mesurés directement.

Le problème de l'estimation d'état a été initialement introduit par Kalman dans son article fondateur de 1960 [Kal60]. En 1961, Kalman a étendu cette étude aux systèmes stochastiques linéaires, fournissant une estimation d'état optimale garantie en présence de bruit gaussien [KB61]. Luenberger a également contribué au domaine en présentant une théorie généralisée des observateurs d'état pour les systèmes déterministes linéaires et en introduisant les concepts d'observateurs d'état réduits et minimaux [Lue64]. Il a de ce fait amélioré la compréhension et l'applicabilité des techniques d'estimation d'état.

La prévalence des systèmes non linéaires a conduit au développement de méthodes d'estimation d'état au-delà des observateurs linéaires. Le Filtre de Kalman Étendu [Kal97] aborde la non-linéarité par une linéarisation, mais sa précision est localement limitée, ce qui peut entraîner des inexactitudes en présence d'un comportement fortement non linéaire.

En contraste avec ces approches, l'estimation à horizon glissant (*MHE - Moving Horizon Estimation*) [MM95] est une méthode d'estimation d'état basée sur l'optimisation. Dans la *MHE*, l'estimation d'état est réalisée en résolvant un problème d'optimisation en temps réel qui minimise une fonction de coût de contrôle optimal approximative sur un horizon glissant défini. Cette méthode présente l'avantage de gérer diverses distributions d'incertitude tout en prenant explicitement en compte les contraintes physiques sur les états. De plus, la *MHE* montre un grand potentiel, en particulier pour l'estimation d'état non linéaire, en raison de ses propriétés de stabilité robuste [Sch+23]. La performance de la *MHE* s'améliore à mesure que la longueur de l'horizon récursif augmente. Cependant, cela entraîne également une charge de calcul proportionnelle à la taille de l'intervalle d'horizon considéré.

Dans le cadre de nos travaux, nous avons proposé une méthodologie pour entraîner un modèle d'apprentissage profond à apprendre une solution au problème d'optimisation de la *MHE*, spécifiquement pour l'estimation complète d'état à partir d'observations partielles et bruitées du système. À la sous-section suivante nous rappelons un état de l'art sur les réseaux de neurones pour l'estimation d'état et la résolution de problèmes inverses. Nous décrivons ensuite notre méthode à la sous-section 3.4.3. Enfin nous réalisons une synthèse des résultats que nous avons obtenus et proposons une discussion.

3.4.2 Réseaux de neurones pour l'estimation d'état et la résolution de problèmes inverses

Récemment, des avancées majeures des techniques d'apprentissage profond ont fait émerger un intérêt croissant pour l'application de ces méthodes dans le domaine des systèmes dynamiques. On pourra citer *KalmanNet* [Rev+22] qui est un estimateur d'état basé sur un réseau de neurones qui apprend à partir de données dans un cadre de filtrage de Kalman. Il traite des dynamiques non linéaires et des informations partielles en incorporant la structure d'un modèle d'espace d'état dans une architecture de réseau de neurones récurrents (*RNN*). *KalmanNet* gère efficacement les non-linéarités et les inadéquations de modèle, surpassant les méthodes de filtrage traditionnelles dans des situations où les connaissances du domaine sont à la fois précises et imprécises. De même, *Backprop KF* [Haa+16] exploite les *RNN* dans un cadre de filtrage d'état probabiliste, permettant la conception d'architectures de réseau spécifiquement adaptées à l'estimation d'état. La méthode est conçue pour entraîner des estimateurs d'état qui utilisent des données complexes issues de capteurs, tels que des images, et montre des améliorations significatives par rapport aux approches génératives standard et aux *RNN* classiques.

Dans le contexte de l'estimation à horizon glissant (*Moving Horizon Estimation - MHE*), [MWZ22] aborde le problème de l'estimation d'état en ligne et de la calibration de paramètres

pour des systèmes linéaires contraints. Ils emploient pour ce faire une couche d'optimisation convexe différentiable pour formuler un estimateur d'état *MHE* en utilisant la descente de gradient stochastique (*SGD - Stochastic Gradient Descent*). Leur méthode est adaptée aux systèmes linéaires contraints avec incertitude paramétrique et montre une amélioration des performances par rapport à la formulation *OMHE*. Dans un contexte étroitement lié, [Cha+23] décrit une approche en deux étapes basée sur la *MHE* et orientée par les données. Plus concrètement, les auteurs exploitent un AutoEncodeur pour construire un modèle de substitution capable de préserver les états de « jumeaux numériques » complexes et exigeants en termes de calcul. Par la suite, l'AutoEncodeur est utilisé pour l'estimation d'état en ligne à partir de mesures bruitées.

On remarque ainsi un changement de paradigme récent dans la résolution des problèmes inverses, où les modèles d'apprentissage automatique sont de plus en plus employés. Ces modèles apprennent à optimiser et à fournir directement des solutions à des problèmes d'optimisation, améliorant de ce fait la précision et l'efficacité computationnelle (coût de calcul) des estimations.

Dans ce contexte, la méthode de *MHE* implique d'estimer l'état en résolvant un problème inverse basé sur une relation implicite entre l'état initial d'un problème à valeurs initiales (*IVP - Initial Value Problem*) et une séquence d'observations du système. Identifier l'état initial nécessite de minimiser cette formulation implicite en utilisant un solveur itératif, ce qui est connu pour être assez chronophage. C'est pourquoi, il y a un intérêt croissant à utiliser des modèles d'apprentissage automatique pour apprendre des solutions approximatives au problème de la *MHE*. Par exemple, dans [Ale+08], un perceptron à une couche est estimé en résolvant un problème de programmation non linéaire afin d'apprendre de manière hors ligne à minimiser un problème d'optimisation *MHE*. Dans [KL21], un perceptron multicouche est utilisé pour apprendre des solutions approximatives aux problèmes de *MHE* et de contrôle prédictif de modèle (*MPC - Model Predictive Control*), en utilisant des données générées par les algorithmes d'*OMHE* et de *MPC*. Ceci leur permet de réduire considérablement le coût de calcul pour obtenir des estimations d'état. En partant dans une direction similaire, la méthode que nous proposons repose sur une formulation entièrement différentiable du critère d'optimisation de la *MHE*, permettant d'apprendre à un modèle d'apprentissage profond à inférer directement la solution de la *MHE* à partir des données de mesure sans s'appuyer sur l'estimation d'*OMHE* comme proposé dans [KL21]. Nous décrivons ainsi notre méthode à la sous-section 3.4.3 suivante.

3.4.3 Méthode proposée

3.4.3.1 Les 3 points clés de notre contribution

Ci-dessous, nous proposons un résumé de notre méthode et des contributions qu'elle apporte :

1. En s'appuyant sur la différentiation automatique (*AD - Automatic Differentiation*), notre méthode repose sur une formulation entièrement différentiable du problème d'optimisation de l'estimation à horizon glissant en ligne (*Online MHE*) [Bay+17]. Cela permet à

un modèle d'apprentissage profond d'apprendre une relation directe qui approxime la solution du problème pendant une période d'entraînement hors ligne sur un ensemble de données contenant des observations du système.

2. Après la phase d'entraînement, le modèle d'apprentissage profond peut fournir une estimation d'état avec un temps de calcul significativement amélioré par rapport à l'algorithme d'*OMHE*, et ce en ne s'appuyant sur aucun processus d'optimisation supplémentaire.
3. Le processus d'entraînement exploite les biais inductifs de l'apprentissage profond pour estimer des solutions qui minimisent globalement le critère d'*OMHE* sur l'ensemble du jeu de données à tous les instants. Ceci permet ainsi d'éliminer la nécessité des termes de coût d'arrivée comme dans l'algorithme d'*OMHE* afin de maintenir la cohérence dans l'estimation séquentielle au fil du temps.

Les résultats expérimentaux que nous avons obtenus ont démontré une performance améliorée par rapport à l'algorithme *OMHE*, même dans des cas complexes, incluant ceux avec des écarts de modèle.

3.4.3.2 Description générale de notre méthode

Notre méthode repose sur une formulation différentiable de bout en bout du problème d'optimisation de l'estimation à horizon glissant (*MHE*), permettant l'entraînement hors ligne d'un modèle d'apprentissage profond pour fournir des estimations d'état qui approximent la solution du critère d'optimisation de la *MHE*.

Le processus d'entraînement de notre méthode suit cinq étapes (voir Figure 3.6) :

1. Le modèle d'apprentissage profond, noté ϕ , reçoit un horizon passé de mesures potentiellement partielles et bruitées, et produit une estimation de l'état initial $x(t_0)$.
2. Une estimation de la trajectoire d'état est obtenue en intégrant numériquement un modèle dynamique connu f , qui représente le système mesuré sous-jacent, en partant de l'estimation de l'état initial $x(t_0)$.
3. L'estimation de la trajectoire d'état est appliquée via la fonction d'observation.
4. La perte (\mathcal{L} - Loss) est calculée pour mesurer la divergence entre la trajectoire finale estimée \hat{y} et les mesures bruitées du système y .
5. Le gradient des paramètres θ du modèle d'apprentissage profond est calculé à partir de la fonction de perte en utilisant un algorithme de différenciation automatique. Cela permet d'optimiser le modèle via la descente de gradient afin de prédire des estimations d'état qui s'alignent le mieux avec les trajectoires mesurées.

Une fois l'entraînement terminé, le modèle d'apprentissage profond apprend à estimer des solutions qui approximent le critère de la *MHE*, sans avoir à supporter les coûts supplémentaires liés à l'intégration numérique du système et aux solveurs en ligne, comme dans la formulation en ligne de l'estimation à horizon glissant (*OMHE*).

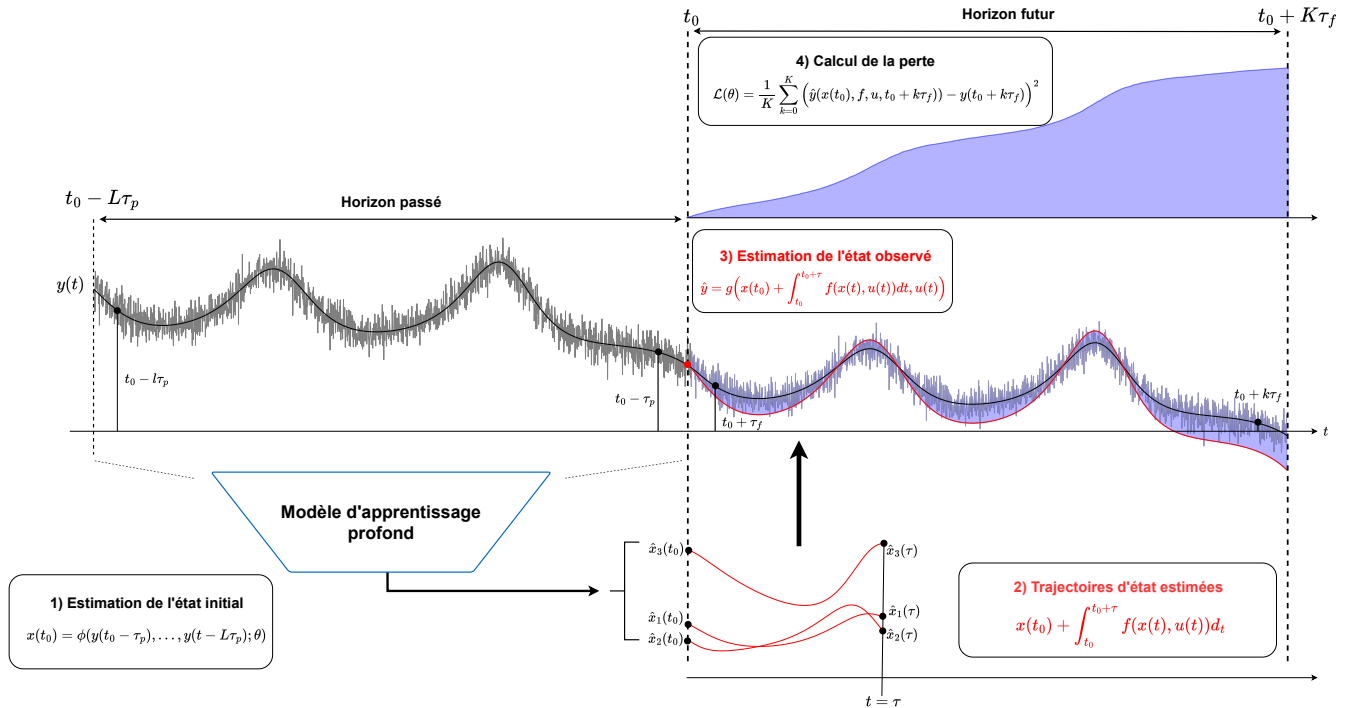


FIGURE 3.6 – Description des étapes du processus d’entraînement

3.4.3.3 Description formelle et détaillée de notre méthode

Nous considérons un système général décrit par une dynamique non linéaire continue de dimension finie de la forme suivante :

$$\dot{x}(t) = f(x(t), u(t)) \quad (3.17)$$

$$y(t) = g(x(t), u(t)) + \eta \quad (3.18)$$

avec l'état $x \in \mathbb{R}^n$, le contrôle $u \in \mathbb{R}^m$, l'observation $y \in \mathbb{R}^p$, et les applications lisses : $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ correspondant aux équations du système dynamique, la fonction de mesure $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$ et η une variable aléatoire correspondant au bruit de mesure. Une stratégie initiale pour l'estimation d'état consiste à simuler les équations f en parallèle avec une gamme d'états initiaux probables choisis a priori, et à écarter progressivement ceux dont la trajectoire de sortie $y(t)$ s'écarte de manière significative de l'observation. Si la sortie $y(t)$ détermine de manière unique une solution de manière asymptotique, l'ensemble des états initiaux réalisables converge vers une seule possibilité. En revanche, si plusieurs états initiaux continuent à fournir des solutions réalisables, cela correspond à un système avec des états non observables.

Nos travaux mettent l'accent sur les systèmes où l'état est globalement observable. Sur cette base, le problème d'estimation d'état peut être formulé comme un problème d'optimisation

avec le critère suivant :

$$\mathcal{L}(x(t_0)) = \int_{t_0}^{t_0+T_f} \left(\hat{y}(x(t_0), f, u, \tau) - y(\tau) \right)^2 d\tau \quad (3.19)$$

$$x^*(t_0) = \min_{x(t_0)} \mathcal{L}(x(t_0)) \quad (3.20)$$

avec x^* correspondant à l'état initial qui minimise la différence entre la trajectoire simulée \hat{y} et la trajectoire mesurée de la variable d'état y . Ce problème de minimisation est résolu sur un horizon de temps $\tau \in [t_0, T_f]$ obtenu à partir de la solution d'un problème à valeurs initiales (IVP) impliquant la dynamique du système f sous l'influence de l'entrée de commande u :

$$\hat{y}(x(t_0), f, u, \tau) = g\left(x(t_0) + \int_{t_0}^{t_0+\tau} f(x(t), u(t)) dt\right) \quad (3.21)$$

La méthode que nous proposons vise à reconstruire toutes les variables d'état à partir de mesures du système qui sont bruitées et partielles y , en supposant que $n > p$, en utilisant un modèle d'apprentissage profond ϕ qui apprend à prédire l'estimation de l'état initial $x(t_0)$, en minimisant \mathcal{L} sur l'horizon futur T_f :

$$x(t_0) = \phi\left(y(t_0), y(t_0 - \tau_p), \dots, y(t_0 - L\tau_p; \theta)\right) \quad (3.22)$$

avec L le nombre total d'échantillons passés considérés, tel que $T_p = t_0 - L\tau_p$ corresponde à l'horizon passé, et θ représentant la paramétrisation du modèle d'apprentissage profond. À des fins d'entraînement, la fonction de coût (3.19) est reformulée dans une version discrétisée :

$$\mathcal{L}(\theta) = \frac{1}{MK} \sum_{m=0}^M \sum_{k=0}^K \left(\hat{y}_m(\phi, f, u, t_0 + k\tau_f) - y_m(t_0 + k\tau_f) \right)^2 \quad (3.23)$$

τ_f représente l'intervalle de temps entre chaque échantillon futur, et l'horizon total de prédiction est donné par $T_f = t_0 + K\tau_f$, où K est le nombre d'échantillons prédits. M correspond au nombre de trajectoires d'entraînement disponibles dans l'ensemble de données \mathcal{D} . Les estimations de trajectoire d'état \hat{y} sont obtenues à l'aide d'une méthode d'intégration numérique :

$$\hat{y}(\phi, f, u, t_0 + k\tau_f) = g \circ F^k \circ F^{k-1} \circ \dots \circ F^0(x(t_0), u(t_0)) \quad (3.24)$$

avec $F(x(t), u(t)) = x(t + \tau_f)$ correspondant à une méthode d'intégration d'Euler à un pas partant de $x(t_0)$:

$$x(t + \tau_f) = f(x(t), u(t))\tau_f + x(t) \quad (3.25)$$

En implémentant les modèles f et g via PyTorch, cela permet de construire un graphe computationnel du schéma d'intégration numérique et, en utilisant l'algorithme AD (Auto-Différentiation), de rendre le problème d'optimisation de la MHE entièrement différentiable afin d'obtenir les gradients par rapport à la paramétrisation θ en rétropropageant les gradients au fil du temps, de

$t_0 + k\tau_f$ à t_0 :

$$\frac{\partial \hat{y}(\phi, f, u, t_0 + k\tau_f)}{\partial \theta} = \frac{\partial g}{\partial F^k} \prod_{i=0}^{k-1} \frac{\partial F^{i+1}}{\partial F^i} \frac{\partial F^0}{\partial \phi} \frac{\partial \phi}{\partial \theta} \quad (3.26)$$

La méthode que nous proposons diffère de l'algorithme d'*OMHE* car il n'y a pas de coût d'arrivée dans la perte (3.23). Dans le contexte séquentiel de l'algorithme d'*OMHE*, ce coût est utilisé pour pénaliser la différence entre l'estimation précédente et l'estimation actuelle. Cela permet de maintenir la cohérence entre les estimations successives et d'augmenter "virtuellement" la longueur de la fenêtre glissante, en réinjectant la connaissance de l'estimation d'état antérieure dans l'estimation actuelle. La méthode que nous proposons transforme le problème d'optimisation localisé à un instant donné dans le contexte de l'*OMHE* en un problème d'optimisation qui s'applique à toutes les étapes temporelles disponibles dans \mathcal{D} . Dans ce cas, l'ajout d'un coût d'arrivée est inutile car le modèle d'apprentissage profond apprend intrinsèquement à préserver la continuité de ses prédictions en minimisant la perte (3.23) sur l'ensemble du jeu de données \mathcal{D} .

De plus, la fonction de perte de la *MHE* proposée, entièrement différentiable, permet de s'émanciper du besoin de superviser l'entraînement directement avec des estimations d'état générées par l'algorithme d'*OMHE*, comme dans [KL21]. Elle étend la méthode à des techniques d'optimisation plus efficaces que dans [Ale+08], telles que la descente de gradient, permettant ainsi l'utilisation de modèles d'apprentissage profond plus complexes.

Une autre de nos motivations à la formulation de la *MHE*, par rapport à d'autres estimateurs d'état, est la capacité d'appliquer des contraintes sur l'estimation d'état en ajoutant des termes de pénalisation au critère d'optimisation. Bien que nous ne l'abordons pas dans nos travaux, l'imposition de contraintes sur l'état estimé peut être facilement proposée en extension de notre méthode en ajoutant des termes à la fonction de perte (3.23).

3.4.4 Synthèse des résultats

Méthodes comparées et cas d'étude. Nous avons évalué les performances de différentes architectures d'apprentissage profond intégrées au sein de notre méthode : *CNN*, *GRU*, *CNN-T*. Nous avons comparé ces performances avec celles de l'algorithme d'*OMHE* comme méthode compétitrice.

Les expériences ont été réalisées sur trois études de cas : Van der Pol, Duffing et sur le modèle de véhicule Dof10.

Nous avons utilisé l'erreur quadratique moyenne (RMSE - root mean square error) comme critère de performance pour évaluer l'ensemble des méthodes comparées. La RMSE a été calculée pour chaque horizon temporel passé T_p et futur T_f ¹.

1. Le détail des implémentations est disponible sur GitHub : https://github.com/N9TT-9G0A-B7FQ-RANC/ICRA_2025.

Résultats. Le détail des résultats expérimentaux est disponible dans notre article soumis en conférence et en cours de relecture à date d'écriture de ce mémoire. Il est accessible au travers de ce lien <https://ngutowski.fr/article2.pdf>.

Nous proposons une synthèse des résultats dans le Tableau 3.1. Ce tableau compare les différentes méthodes en termes d'avantages, d'inconvénients, de temps de calcul et de RMSE moyenne obtenue.

Méthode	Avantages	Inconvénients	RMSE Moyenne obtenue	Temps de Calcul (en ms)
<i>Online Moving Horizon Estimation (OMHE)</i>	Robustesse dans les systèmes non linéaires, gestion des incertitudes	Coût de calcul élevé, sensible à la taille de l'horizon	Varié selon l'horizon, tendance à augmenter	Varié, généralement élevé
<i>Convolutional Neural Network (CNN)</i>	Bonne capture des motifs locaux et dépendances temporelles	Performance variable selon la complexité des motifs, moins consistant	Modéré, dépend des motifs locaux capturés	Faible, constant
<i>Gated Recurrent Unit (GRU)</i>	Excellente capture des dépendances séquentielles	Biais inductif bien adapté, mais coûteux en calcul pour les longues séquences	Faible, meilleure performance globale	Modéré, constant
<i>CNN-Transformer (CNN-T)</i>	Capacité de capturer des dépendances complexes	Perte partielle d'information temporelle, dégradation des performances	Modéré à élevé, performance moins consistante	Modéré, constant

TABLE 3.1 – Comparaison des méthodes expérimentées avec avantages, inconvénients, RMSE et temps de calcul.

Plus précisément, les *GRU* ont obtenu les meilleures performances de RMSE sur l'ensemble des études de cas, suivis de près par les *CNN*. Cela est dû à la capacité des *GRU* à capturer de manière efficace les dépendances séquentielles, essentielle pour les tâches d'estimation d'état dans des systèmes dynamiques. Les *CNN*, bien que légèrement inférieurs en termes de performance, sont une alternative tout à fait viable dans certaines situations. En revanche, les modèles *CNN-T* ont obtenu des performances significativement dégradées, potentiellement dues à une perte partielle d'informations temporelles dans leur représentation interne.

Évolution des performances en fonction des horizons temporels passés et du temps d'intégration. Nous avons identifié deux schémas distincts dans l'évolution des performances, mesurées par la RMSE, en fonction des horizons temporels passés (T_p) et du temps d'intégra-

tion (T_f) (voir Tableau 3.2), pour notre méthode utilisant l'apprentissage profond et l'algorithme *OMHE* :

1. **Schéma 1 : Van der Pol et Duffing.** Dans ces études de cas, il n'y a pas de désaccord entre le modèle utilisé pour l'entraînement et celui pour la simulation. Les deux approches (notre méthode et *OMHE*) sont pertinentes, mais notre méthode surpasse l'*OMHE* dans certaines configurations :
 - **Performance** : L'augmentation de T_p et T_f améliore les performances en réduisant la RMSE pour les deux méthodes, mais notre méthode montre une meilleure réduction de la RMSE.
 - **Convergence** : Au-delà d'une certaine valeur de T_f , l'augmentation du temps d'intégration n'apporte plus d'amélioration significative pour notre méthode. L'*OMHE* quant à lui continue à souffrir d'une légère augmentation de la RMSE.
 - **Interprétation** :
 - (a) Notre méthode bénéficie de l'absence de divergence entre le modèle et les données, ce qui lui permet d'apprendre une représentation suffisamment riche pour estimer les variables d'état non observées plus efficacement que l'*OMHE*.
 - (b) Des temps d'intégration plus longs permettent à notre méthode d'optimiser la perte de la *MHE* sur des trajectoires plus longues, rendant le gradient plus robuste aux fluctuations de bruit.
 - (c) L'augmentation de T_f réduit la variance de notre méthode, mais a un effet moindre sur l'*OMHE*.
 - (d) Pour des valeurs intermédiaires de T_p , notre méthode montre une meilleure performance avant d'atteindre un point de stagnation plus tardif par rapport à l'*OMHE*.
2. **Schéma 2 : Étude de cas Dof10.** Dans cette étude de cas, il existe un décalage entre le modèle d'entraînement et celui utilisé pour simuler les trajectoires. Notre méthode montre une meilleure robustesse comparée à l'*OMHE* :
 - **Performance** : L'augmentation de T_f provoque une dégradation significative des performances pour l'algorithme *OMHE* avec une augmentation de la RMSE. En revanche, notre méthode, bien qu'affectée, reste relativement stable.
 - **Interprétation** :
 - (a) Le modèle d'apprentissage profond apprend à estimer les états initiaux en minimisant la différence entre les trajectoires biaisées et celles observées, et il est plus robuste que l'*OMHE* face à ce biais.
 - (b) Ce biais empêche le modèle d'*OMHE* de tirer parti des horizons temporels plus longs (T_p), tandis que notre méthode reste plus stable face à l'augmentation de T_f .

Résumé des résultats. En conclusion, notre méthode d'apprentissage profond surpasse systématiquement l'algorithme *OMHE* en termes de réduction de la RMSE et de stabilité, no-

Études de cas	Effet de l'augmentation de T_p	Effet de l'augmentation de T_f
Van der Pol & Duffing	Amélioration des performances pour notre méthode et l'OMHE, avec notre méthode plus performante	Convergence après une certaine valeur pour notre méthode, tandis que l'OMHE continue de se dégrader
Dof10	Dégradation significative des performances pour l'OMHE, notre méthode reste plus stable	Augmentation de la RMSE pour l'OMHE, alors que notre méthode conserve une stabilité relative

TABLE 3.2 – Résumé des effets de l'augmentation des paramètres temporels sur les performances.

tamment dans des cas complexes comme l'étude de cas Dof10 où l'OMHE est particulièrement affectée par le biais du modèle. Ces résultats montrent clairement l'avantage de notre approche, qui offre une meilleure robustesse tout en réduisant le temps de calcul.

3.4.5 Discussion

Nous avons présenté une approche novatrice pour l'estimation d'état dans des systèmes dynamiques non linéaires en utilisant des modèles d'apprentissage profond entraînés sur une formulation entièrement différentiable du problème d'optimisation de la MHE (Estimation à Horizon Glissant). L'objectif principal a été de fournir des estimations d'état précises à partir de mesures bruitées et partielles, tout en réduisant les coûts computationnels associés aux méthodes traditionnelles, telles que l'OMHE (Estimation à Horizon Glissant en Ligne). Les résultats expérimentaux ont montré que notre méthode non seulement améliore la précision des estimations dans divers scénarios, mais réduit aussi de manière significative le temps de calcul, la rendant particulièrement adaptée aux applications en temps réel.

Notre approche met en lumière les avantages indéniables de l'intégration de l'apprentissage profond dans les problèmes d'estimation d'état. Nous exploitons ainsi les biais inductifs d'architectures comme les GRU et les CNN, permettant de capturer les dépendances temporelles et d'apprendre les représentations des états non observés. Le cadre différentiable proposé permet également une meilleure flexibilité, en éliminant la nécessité des coûts d'arrivée spécifiques, et permet l'optimisation du critère de la MHE sur l'ensemble des données d'entraînement. Cette approche permet de surmonter certaines limites de l'OMHE, notamment dans les cas où il existe des écarts entre le modèle utilisé pour l'entraînement et celui employé pour la simulation des données, comme illustré dans l'étude du modèle Dof10.

Cependant, certaines limites doivent être soulignées. Premièrement, la méthode repose sur l'hypothèse d'une grande quantité de données d'entraînement pour garantir une bonne généralisation sur tout l'espace des états et des commandes. Dans des scénarios où les données sont rares ou coûteuses à obtenir, la performance du modèle peut en souffrir. De plus, cette approche fournit des estimations déterministes, sans tenir compte des incertitudes associées aux

états estimés. Bien que cela soit acceptable pour certaines applications, d'autres contextes, tels que la robotique ou les systèmes critiques, nécessitent des informations probabilistes afin de quantifier ces incertitudes.

Un autre point à prendre en compte est la dépendance au modèle a priori lors de la phase d'entraînement. Lorsqu'il y a divergence entre le modèle utilisé pour l'entraînement et la dynamique réelle du système, comme dans l'étude du Dof10, le biais introduit peut affecter la qualité des prédictions. Cependant, la méthode que nous avons proposée a montré une robustesse face à ces divergences même si des améliorations peuvent être envisagées (p.ex., intégration de modèles d'apprentissage reposant entièrement sur les données pour réduire la dépendance aux modèles physiques possiblement biaisés).

En conclusion, notre travail constitue une avancée importante dans l'application des techniques d'apprentissage profond à l'estimation d'état dans des systèmes dynamiques complexes. Il démontre que les architectures modernes d'apprentissage profond, combinées à une formulation différentiable de la *MHE*, peuvent fournir des estimations d'état précises tout en réduisant significativement les temps de calcul. Bien que des problématiques subsistent, notamment en termes de généralisation et de traitement des incertitudes, la flexibilité et les performances démontrées par cette méthode ouvrent des perspectives prometteuses pour son application à d'autres contextes d'estimation dans les systèmes dynamiques, tels que l'estimation de paramètres ou la prédiction de trajectoires.

3.5 Réflexion critique

Il est pertinent de réfléchir aux implications éthiques des méthodes et technologies que nous proposons dans le contexte plus spécifique de la modélisation mathématique et de l'automatisation.

Entre modélisation et réalité. Nos travaux reposent sur des modèles mathématiques permettant d'estimer les états et les paramètres de systèmes physiques, comme la dynamique des véhicules. Cependant, comment ces modèles mathématiques représentent-ils le monde réel ?

- **Modèles et simplifications** : Les modèles employés (p.ex., le modèle de bicyclette dans les articles sur les dynamiques latérales) sont souvent des simplifications qui s'émancipent de certains aspects complexes du système. Dans quelle mesure un modèle peut refléter fidèlement la réalité suite à une simplification trop importante ? En effet, les simplifications facilitent le calcul, mais elles introduisent également de possibles biais. Ces biais peuvent réduire l'applicabilité du modèle dans des cas plus complexes du monde réel.
- **Imprévus et complexité** : Certains modèles échouent à capturer des dynamiques imprévues ou des interactions complexes tant il reste difficile de modéliser la réalité. Cette question montre les limites de la modélisation en général, et la nécessité d'utiliser des

méthodes comme l'inférence bayésienne ou l'apprentissage par transfert pour combler les lacunes entre le modèle et les données empiriques.

Éthique et responsabilité dans l'automatisation. L'automatisation croissante des processus d'estimation, notamment via des algorithmes d'apprentissage profond ou des techniques bayésiennes, soulève plusieurs questions éthiques.

- **Transparence des algorithmes** : Les systèmes décrits dans nos travaux fonctionnent souvent comme des « boîtes noires ». La transparence et l'explicabilité de ces algorithmes sont essentielles, surtout dans des applications critiques (comme la dynamique des véhicules), pour garantir que les décisions prises par ces systèmes puissent être comprises et vérifiées.
- **Responsabilité** : Qui peut/doit être tenu pour responsable lorsqu'une mauvaise décision est prise par un système automatisé ? Cette question se pose dans le cadre de l'évaluation des véhicules mais peut tout aussi bien s'appliquer dans le cadre des décisions médicales (voir Chapitre 2). Le responsable est-il le concepteur du modèle ? celui qui l'a entraîné ou l'utilisateur final ? Ces questions sont incontournables avant de pouvoir déployer ces solutions dans des systèmes où les décisions pourraient être prises en temps réel ou sans intervention humaine.

En conclusion. Nos travaux de recherche s'inscrivent dans une tendance à l'automatisation croissante et à l'utilisation d'algorithmes avancés dans des systèmes complexes. La philosophie des sciences et l'éthique offrent des cadres essentiels pour réfléchir aux limites, à l'interprétation et aux responsabilités liées à l'utilisation de ces technologies.

OPTIMISATION DES REPRÉSENTATIONS SYMBOLIQUES MUSICALES POUR L'APPRENTISSAGE PROFOND

Symbolic Music generation potential...

« *DeepBach, a probabilistic model together with a sampling method which is flexible, efficient and provides musically convincing results even to the ears of professionals.* [HPN17]

G. Hadjeres, F. Pachet & F. Nielsen - 2017 »

Sommaire

4.1 Concepts généraux sur la génération de musique symbolique par apprentissage profond	149
4.2 Positionnement de nos travaux de recherche	159
4.3 Contribution logicielle : <i>Miditok</i>	160
4.4 Impact de la tokenisation du temps et de la durée des notes sur la modélisation de la musique symbolique	162
4.5 Byte Pair Encoding pour la musique symbolique	167
4.6 Réflexion critique générale sur les <i>LLMs</i>	171

Ce chapitre a pour objectif

- de rappeler les principes et les enjeux en génération de musique symbolique, en commençant par les techniques d'apprentissage profond employées aujourd'hui et en mettant progressivement l'accent sur les problématiques de tokenisation ;
- de réaliser la synthèse de nos contributions sur la tokenisation de la musique symbolique ;
- de mettre en œuvre in fine une discussion qui portera à la fois sur les travaux existants dans la littérature et sur nos propres contributions à ce champ de recherche.

Contexte des contributions de ce chapitre

Historique. Ce chapitre et les travaux qui en sont issus font référence à la thèse CIFRE de Nathan Fradet [Fra24] dirigée par Jean-Pierre Briot (LIP 6) et Amal El-Fallah (LIP 6) et que j'ai co-encadré de 2021 à 2024. Cette thèse était intitulée : « Apprentissage automatique pour la modélisation de musique symbolique ». Cette thèse a émergé à la suite de mon encadrement du projet de fin d'études de Nathan Fradet chez Aubay, qui portait notamment sur l'apprentissage profond pour la génération de musique. Pour répondre à son souhait de réaliser une thèse industrielle, j'ai co-construit le sujet en collaboration avec le futur doctorant, l'entreprise, le LIP 6 et l'ESEO.

Équipe. L'équipe d'encadrement constituée était tripartite : le LIP 6 pour la direction de thèse, l'ESEO, le LERIA pour le co-encadrement, et l'entreprise Aubay dans le cadre de la CIFRE.

Environnement et organisation. Les travaux de thèse se sont déroulés pour le doctorant entre Angers et Paris impliquant un travail de recherche à la fois en distanciel et en présentiel. Concernant la thèse les réunions de suivi se déroulaient au fil de l'eau mais nous avons un point récurrent systématique au moins une fois par mois.

Mon implication et mes contributions. J'ai été impliqué dans le co-encadrement à hauteur de 30%. Cette mission englobait :

- le suivi régulier du doctorant ;
- le conseil scientifique et stratégique ;
- la participation à l'écriture et la revue d'articles ;
- la participation aux conférences dans lesquelles nous avons publié.

Note importante. Les sections de ce chapitre s'appuient sur les travaux suivants : [Fra24], [Fra+23a], [Fra+23b] et [Fra+21].

Ces travaux ont été réalisés en collaboration avec les co-auteurs et les contributions intellectuelles sont ainsi partagées entre eux. De même, le mémoire de thèse de Nathan Fradet résulte d'un travail conjoint entre le doctorant (producteur du mémoire) et ses co-encadrants (relecteurs et conseillers). Ainsi, certaines sections de ce chapitre peuvent provenir de traductions, de reformulations ou encore de formulations textuelles de ces travaux originaux. Les articles mentionnés ci-dessus sont disponibles dans les annexes.

4.1 Concepts généraux sur la génération de musique symbolique par apprentissage profond

Cette section rappelle un rapide état de l'art sur la génération de musique symbolique par apprentissage profond. Pour ce faire, elle repose sur des éléments présentés dans [Fra24 ; Fra+21 ; Fra+23a ; Fra+23b].

4.1.1 Rappel sur les représentations de la musique.

La musique peut être représentée sous deux formes principales [Roa96] : la représentation symbolique et la représentation audio.

La première se concentre sur les aspects théoriques de la musique [AD01], tels que les notes, les silences, les rythmes et les tempos, souvent matérialisés par des partitions. Elle est particulièrement utile pour la composition et l'organisation des structures musicales, permettant aux musiciens de lire et écrire la musique en se basant sur la théorie.

La représentation audio, quant à elle, correspond à la forme physique du son lors de son interprétation. Elle est représentée par des signaux acoustiques qui peuvent être visualisés sous forme d'onde ou de spectrogrammes. Ce format est essentiel pour capturer les caractéristiques sonores comme le timbre, l'intensité et la texture, aspects clés dans la production de genres tels que la musique électronique, la pop ou le rap, où la création sonore et le traitement des sons jouent un rôle central.

Ces deux formes pourtant différentes sont souvent utilisées conjointement dans les processus modernes de créations. Les avancées dans les logiciels de production musicale (DAW) ont notamment facilité l'intégration de ces deux formats. Ceci permet ainsi aux musiciens et producteurs de composer des mélodies via des représentations symboliques tout en manipulant des sons audio pour enrichir la texture sonore. Cela permet aussi une approche plus complète de la création musicale en associant la précision théorique et la richesse sonore.

Nos travaux ayant spécifiquement ciblé la génération de la musique symbolique, ce chapitre portera principalement sur cette thématique.

4.1.2 Formats de données pour la représentation et la génération de musique symbolique

4.1.2.1 Formats numériques

La génération de musique symbolique repose sur divers formats numériques qui permettent de représenter des éléments musicaux abstraits tels que les notes, le rythme et les silences. Ces formats, largement utilisés pour l'échange et l'analyse de données musicales, sont essentiels dans les systèmes de génération assistée par ordinateur.

Nous rappelons ci-dessous ces différents formats numériques :

1. **Le format MIDI (*Musical Instrument Digital Interface*)**. Il est standardisé en 1983, permet la communication entre instruments numériques et ordinateurs. Le protocole *MIDI* repose sur une série d'événements structurés en canaux, chaque événement étant composé d'un temps relatif (en *ticks*) et d'un message qui précise l'action musicale (*Note On*, *Note Off*, changement de programme, etc.). Les spécifications *GM1* et *GM2*¹ garantissent l'interopérabilité des dispositifs *MIDI* entre fabricants. Le format *MIDI* permet également la gestion simultanée de plusieurs instruments, notamment grâce aux 16 canaux disponibles, avec le canal 10 réservé aux percussions.

1. Spécifications GM : <https://www.midi.org/specifications>

2. **Le format symbolique *ABC notation*.** Il est utilisé principalement pour les mélodies monophoniques issues des traditions folkloriques européennes. Ce format simple se base sur des lettres représentant les notes et des symboles indiquant la durée des notes par rapport à une longueur par défaut.
3. **Le format *MusicXML (format ouvert), inspiré du XML*.** Il permet de structurer des partitions musicales de manière hiérarchique. Chaque mesure, note et attribut sont décrits dans un format lisible et inter-opérable entre différents logiciels de notation musicale. *MusicXML* est adapté pour des compositions complexes et polyphoniques.

Ces formats symboliques, de *MIDI* à *MusicXML*, sont des outils incontournables pour la représentation de la musique sous une forme « manipulable » par des systèmes informatiques. Ils se constituent ainsi comme une base pour les données d'apprentissage dans les modèles de génération musicale reposant sur l'Intelligence Artificielle.

Nos travaux ont plus particulièrement reposé sur l'usage du format *MIDI*. La figure 4.1 montre à quoi peut ressembler une séquence de messages MIDI. Tous les messages et leurs significations sont décrits dans les spécifications GM1 et GM2. Pour nos besoins, les messages les plus importants et pertinents, que nous avons analysé pour extraire les informations musicales, sont :

- **Note On** : une note est jouée. Ce message spécifie la hauteur de la note (*pitch*) et sa vitesse (*velocity*) ;
- **Note Off** : une note est relâchée. Ce message indique la note à arrêter et la vitesse de la note ;
- **Program Change** : spécifie l'instrument à utiliser ;
- **Control Change** : un paramètre de contrôle est modifié ou appliqué. La molette de modulation, la pédale forte, ou le contrôle du volume sont, par exemple, des effets transcrits sous forme de messages *Control Change*.

Dans la figure 4.1 par exemple : une trompette est utilisée et la note Do (*pitch* 60, c.-à-d. C4 ou Do central de 261,626Hz) est jouée (*Note On*) à une intensité modérée (*Velocity* 50 sur 127).

```
11000101 - 00111001
Program Change, Channel 5 - Instrument 57 (Trumpet)

10010101 - 00111100 - 00110010
Note On, Channel 5 - Pitch 60 - Velocity 50

10000101 - 00111100 - 00000000
Note Off, Channel 5 - Pitch 60 - Velocity 0
```

FIGURE 4.1 – Exemples de messages MIDI : valeurs binaires et leurs significations [Fra24].

4.1.2.2 Cas spécifiques de la représentation sous forme de tokens

Dans le cadre de nos travaux, les modèles d'apprentissage que nous avons employés, nécessitaient une représentation de la musique symbolique, sous forme de tokens. Le processus transformant les données musicales en tokens est connu sous le nom de tokenisation. Dans notre cas, la tokenisation a consisté à transformer les fichiers *MIDI*, en séquences de tokens discrets, compréhensibles et manipulables par les modèles.

La Figure 4.2 illustre plusieurs exemples de séquence de tokens par rapport à deux méthodes de tokenisation issues de la littérature. Ces séquences sont fournies dans une représentation symbolique de la musique, basée sur un encodage *MIDI*.

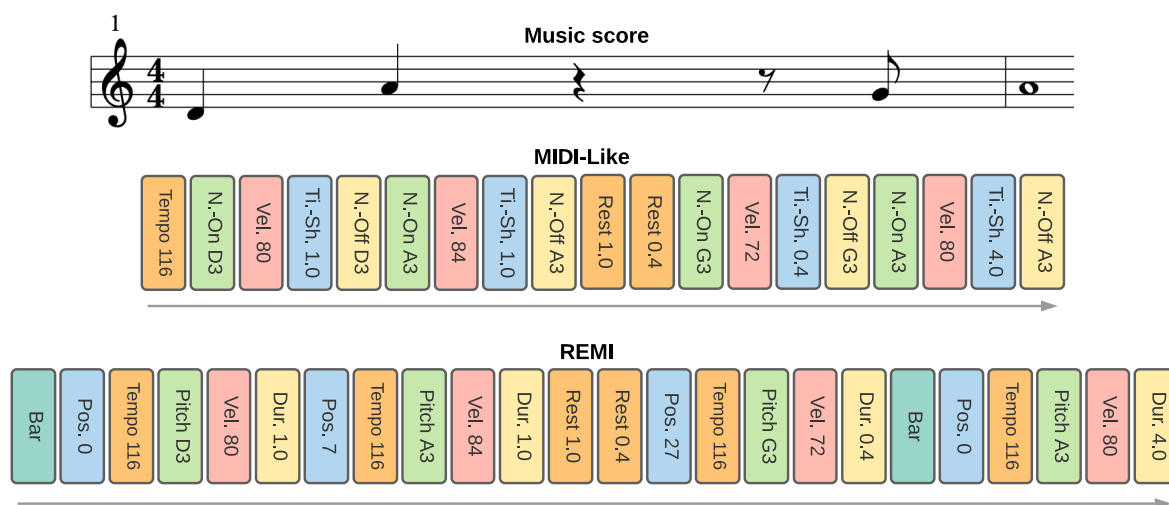


FIGURE 4.2 – Partition musicale et séquences de tokens dans une représentation symbolique de la musique [Fra+21]

Voici quelques exemples de *tokens*, ainsi que leur signification :

- **Bar** : Indique le début d'une mesure ;
- **Pos. (Position)** : Indique la position dans la mesure. Par exemple, "*Pos. 0*" représente le début de la mesure, et "*Pos. 7*" représente une position plus avancée ;
- **Tempo** : Représente le tempo de la pièce en battements par minute (BPM). Ici, il est fixé à 116 BPM ;
- **Pitch** : Indique la hauteur des notes. Par exemple, "D3", "A3" et "G3" représentent les hauteurs musicales (Ré3, La3, Sol3) ;
- **Vel. (Vélocité)** : Indique la dynamique ou l'intensité d'une note, par exemple "*Vel. 80*", "*Vel. 84*", ou "*Vel. 72*" ;
- **Dur. (Durée)** : Indique la durée des notes ou des silences. Par exemple, "*Dur. 1.0*" signifie une durée d'un temps complet, alors que "*Dur. 0.4*" indique une durée plus courte ;
- **Rest (Silence)** : Représente un silence, avec sa durée spécifiée (par exemple, "*Rest 1.0*" pour un silence d'un temps complet et "*Rest 0.4*" pour un silence plus court) ;

- **Time Shift** : indique combien de temps s'écoule avant qu'un autre événement se produise, comme une nouvelle note.

Les différents *tokens* sont alignés dans une séquence temporelle qui décrit les événements musicaux de manière ordonnée.

À la sous-section suivante, nous détaillons plus en détail le principe de tokenisation de la musique symbolique et en dressons les enjeux.

4.1.3 Tokenisation de la musique symbolique

4.1.3.1 Principe de la tokenisation

La tokenisation est une étape incontournable plus généralement dans les méthodes générative employant des architectures de type *Transformers* et bien entendu spécifiquement dans le traitement de la musique symbolique pour l'apprentissage profond. Dans notre cas, la tokenisation consiste à convertir des données musicales, comme p.ex., des fichiers *MIDI*, en séquences de tokens discrets que les modèles peuvent comprendre et manipuler. Ce processus est primordial pour utiliser des modèles de type *Transformers* ou d'autres architectures d'apprentissage profond (*Deep Learning*) qui travaillent sur des séquences de symboles (ou tout simplement de texte dans le cadre de tâches de *NLP*).

4.1.3.2 Principales méthodes de tokenisation

Nous présentons ci-dessous les méthodes de tokenisation les plus connues et renvoyons aux articles associés pour plus de détails.

- **MIDI-Like** [Oor+18] : représente les messages *MIDI* sous forme de tokens. Les notes sont représentées avec les tokens `NoteOn` et `NoteOff`, indiquant leur début et leur fin, et le temps est représenté avec des tokens `TimeShift` ;
- **REMI** [HY20] : signifiant *REvamped MIDI*, il représente la durée des notes avec des tokens explicites `Duration` à la place des tokens de fin `NoteOff`, et le temps est représenté par une combinaison des tokens `Bar` et `Position` indiquant respectivement le début d'une nouvelle mesure et la position du temps dans la mesure actuelle ;
- **REMI+** [Rüt+23] : est une extension de *REMI*, permettant également de représenter les instruments des notes et la signature rythmique ;
- **Structured** [HC21] : similaire à *MIDI-Like*, sauf qu'il représente explicitement la durée des notes comme *REMI* et utilise toujours le même schéma de succession des types de tokens : `Pitch`, `Velocity`, `Duration` et `TimeShift` ;
- **TSD** [Fra+23a] : signifiant *TimeShift & Duration*, il est identique à *MIDI-Like* mais avec des tokens explicites `Duration` ;
- **Compound Word** [Hsi+21] : similaire à *REMI*, mais regroupe plusieurs catégories d'embeddings de tokens afin de réduire la longueur de la séquence pour le modèle. Par exemple, les embeddings de `Pitch`, `Velocity` et `Duration` d'une note sont d'abord concaténés puis projetés pour obtenir un embedding fusionné, et plusieurs couches de

sortie sont utilisées pour prédire ces attributs simultanément. Nous qualifions cette tokenisation de *multi-vocabulaire*, car elle repose essentiellement sur plusieurs vocabulaires distincts ;

- **Octuple** [Zen+21] : également une tokenisation *multi-vocabulaire*, elle fonctionne en fusionnant les embeddings des attributs de chaque note, avec les embeddings `Bar_n` et `Position_p` représentant sa position dans le temps, ce qui donne une séquence de tokens aussi longue que le nombre de notes tokenisées ;
- **MuMIDI** [Ren+20] : une tokenisation *multi-vocabulaire* similaire à *Compound Word*, mais représentant également les programmes des notes (instruments) avec un mécanisme d'encodage positionnel intégré basé sur le nombre de mesures et de positions écoutées ;
- **MMM** [EP20] : une tokenisation pour la génération de musique multi-piste, y compris l'inpainting.

4.1.3.3 Enjeux de la tokenisation en musique symbolique

La tokenisation de la musique symbolique est un processus clé pour la génération musicale à l'aide de modèles d'apprentissage profond. Cependant, elle présente plusieurs problématiques, notamment en ce qui concerne la réutilisation du code et des modèles, la longueur des séquences de tokens, et la nature des informations portées par les tokens. Nous revenons sur ces points importants ci-dessous.

Réutilisation du code et des modèles. Bien que de nombreux travaux partagent leur code source [Hsi+21 ; HY20 ; Zen+21], leur réutilisation reste très souvent difficile. Les méthodes de tokenisation peuvent varier d'un projet à l'autre, avec des dépendances logicielles spécifiques et des pré-traitements différents pour les fichiers *MIDI*. Contrairement à d'autres domaines comme le *NLP*, où des plateformes comme Hugging Face [Jai22] standardisent le partage des modèles et des tokeniseurs, la musique symbolique ne bénéficie pas de tels standards. Il devient alors compliqué de réutiliser des modèles préentraînés et cela peut empêcher de comparer équitablement différentes méthodes de tokenisation.

Contrairement au texte, la musique symbolique peut être représentée de diverses manières en fonction des attributs musicaux à capturer (durée des notes, instruments, tempo). Par exemple, des méthodes comme *MIDI-Like* [Oor+18], *REMI* [HY20] et *CPWord* [Hsi+21] proposent chacune des approches différentes pour encoder la musique en séquences de tokens. Cependant, il reste difficile de déterminer quelle méthode est la plus adaptée pour les tâches spécifiques (génération musicale, classification, etc.), et la plupart des travaux se concentrent uniquement sur la génération musicale.

Longueur des séquences. Un autre problème majeur est la longueur des séquences de tokens. Lorsqu'on représente chaque attribut d'une note (hauteur, vitesse, durée) par un token

distinct, la séquence générée peut devenir très longue, surtout pour des compositions polyphoniques ou avec plusieurs instruments. Cela représente une problématique pour les modèles *Transformers*, dont la complexité croît quadratiquement avec la longueur des séquences. Des méthodes comme *CPWord* [Hsi+21] et *Octuple* [Zen+21] ont cherché à réduire la longueur des séquences en fusionnant certains attributs, mais cela entraîne des contraintes techniques supplémentaires et des résultats potentiellement instables lors de la génération.

Nature des informations portées par les tokens. Les vocabulaires actuels utilisés pour la modélisation de la musique symbolique sont relativement « petits », puisqu'ils vont de 200 à 500 tokens. Cela se situe bien en dessous des vocabulaires typiques dans le *NLP* qui comptent des dizaines de milliers de tokens. Cette contrainte limite l'exploration du potentiel des espaces d'embedding et la capacité des modèles à capturer des représentations complexes. Il est donc important de trouver un équilibre entre la taille du vocabulaire et le nombre de dimensions des embeddings afin de pouvoir améliorer la performance des modèles.

En conclusion, une standardisation des méthodes de tokenisation pour la musique symbolique est nécessaire pour faciliter la réutilisation des modèles préentraînés, réduire la complexité des séquences de tokens, et optimiser les représentations portées par les tokens. Cette standardisation contribuerait à l'avancement de la recherche et des applications pratiques dans ce domaine, notamment en offrant des solutions plus efficaces pour des tâches telles que la génération musicale ou la classification. Elle fût l'objet de notre contribution logicielle *Miditok* que nous décrivons en Section 4.3.

4.1.4 Modèles discrets et continus pour la génération de musique symbolique

La génération de musique symbolique repose principalement sur deux types de modèles : les modèles discrets et les modèles continus.

Les modèles discrets, comme les *RNN* ou les *Transformers*, traitent des séquences d'éléments discrets et sont adaptés à la génération musicale reposant sur des événements musicaux. Parmi les premières implémentations de modèles discrets on peut relever *FolkRNN* [SSK15] et *DeepBach* [HPN17] qui ont montré l'efficacité des *RNN* pour générer des mélodies et chorales. Des encodages spécifiques comme *MIDI-Like* [Oor+18] et *REMI* [HY20] ont quant à eux été proposés pour mieux représenter les événements musicaux et résoudre des problèmes liés à la durée des notes ou à la structuration du temps. Enfin, pour palier les limites liées aux séquences trop longues, des représentations comme *Compound Word* [Hsi+21] et *Octuple* [Zen+21] ont été créées pour permettre de regrouper les tokens associés, réduisant ainsi la longueur des séquences tout en conservant la qualité des informations musicales.

Du côté des modèles continus, la représentation symbolique de la musique sous forme de pianoroll, bien qu'intuitive, présente des limitations dans la gestion des notes maintenues et répétées, comme observé dans *MuseGAN* [Don+18] et *Coconet* [Hua+17]. D'autres approches, telles que *MusicVAE* [Rob+18], intègrent des espaces latents continus pour améliorer la génération musicale, bien que les modèles continus restent moins courants dans ce domaine.

Bien que les modèles continus puissent parfois apporter des résultats intéressants, la majorité des recherches sur la génération de musique symbolique repose sur des modèles discrets, mieux adaptés aux séquences structurées d'événements musicaux. Nos travaux se sont ainsi focalisés sur les modèles discrets. Nous justifions ce choix en comparant les deux approches dans le **Tableau 4.1**.

Caractéristiques	Modèles discrets	Modèles continus
Représentation des données	Séquences d'événements discrets (ex. : <i>Note On/Off</i> , changement de programme)	Matrices ou transformations continues (ex. : espace latent, Pianoroll ¹)
Traitement temporel	Gestion explicite du temps, chaque événement est associé à un moment précis	Temps implicite ou représenté par des blocs continus
Exemples de modèles	<i>FolkRNN, DeepBach, Music Transformer</i>	<i>MuseGAN, Coconet, MusicVAE</i>
Exemples d'encodages	<i>MIDI-Like, REMI, Compound Word, Octuple</i>	Pianoroll ¹ , espace latent
Avantages	Idéal pour modéliser des séquences structurées d'événements musicaux (mélodies, harmonies)	Manipulation flexible, génération de musique multi-pistes, possibilité de synthèse sonore
Inconvénients	Difficulté avec des séquences longues, accumulation d'erreurs lors de la génération	Incohérence dans l'échantillonnage, difficulté à gérer des événements distincts comme les notes maintenues ou répétées
Applications	Compositions musicales avec événements distincts, musique symbolique traditionnelle (<i>MIDI</i> , partitions)	Musique multi-pistes, musique électronique, synthèse sonore

TABLE 4.1 – Différences entre modèles discrets et continus pour la génération de musique symbolique.

4.1.5 Mécanisme d'attention et *Transformers*

Nos travaux ont reposé sur l'utilisation des modèles tels que *GPT-2* et *BERT*. Leur composant essentiel repose sur les *Transformers*. Cette sous-section en présente donc les principes fondamentaux.

L'architecture *Transformer* a été introduite par Vaswani et al. [Vas+17]. Elle est considérée comme une véritable révolution dans le domaine de l'apprentissage profond, d'abord en traitement du langage naturel (*TALN, Natural Language Processing - NLP*), puis dans d'autres domaines comme la vision par ordinateur et dans notre cas la musique symbolique. Contrairement aux architectures séquentielles telles que les *RNNs* (*Recurrent Neural Network*, réseau de neurones récurrents) ou les *LSTMs* (*Long Short-Term Memory*, cellule de longue mémoire à court terme), les *Transformers* reposent uniquement sur le mécanisme d'attention. Ceci élimine les dépendances temporelles et permet un traitement parallèle plus efficace.

1. Le pianoroll n'est pas exactement une représentation continue, et son application à des opérations continues ne garantit pas des résultats très performants. Certains autres modèles plus abstraits (comme des espaces latents) approchent davantage une représentation véritablement continue.

4.1.5.1 Mécanisme d'attention

L'attention permet de modéliser les relations entre les éléments d'une séquence en apprenant des poids qui reflètent l'importance relative de chaque élément par rapport aux autres. Plus précisément, les "clés" (\mathbf{K}) jouent un rôle central en représentant chaque élément de la séquence sous une forme compressée qui capture ses informations contextuelles. Les "requêtes" (\mathbf{Q}) interrogent ces clés pour déterminer les éléments pertinents de la séquence, tandis que les "valeurs" (\mathbf{V}) contiennent les informations finales extraites en fonction des poids calculés. Le mécanisme d'attention par produit scalaire normalisé, utilisé dans les *Transformers*, est défini par :

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax} \left(\frac{\mathbf{Q}\mathbf{K}^\top}{\sqrt{d_k}} \right) \mathbf{V} \quad (4.1)$$

où \mathbf{Q} , \mathbf{K} et \mathbf{V} sont respectivement les matrices des requêtes, des clés et des valeurs dérivées des séquences d'entrée, et d_k est la dimension des clés. La normalisation par $\frac{1}{\sqrt{d_k}}$ permet de stabiliser le calcul des gradients pour des séquences de grande taille.

La **multi-head attention** (attention multi-tête) améliore ce mécanisme en calculant plusieurs mécanismes d'attention en parallèle. Chaque tête d'attention apprend à se concentrer sur différentes parties de la séquence ou à modéliser différents types de relations :

$$\text{MultiHead}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{Concat}(\text{head}_1, \dots, \text{head}_h) \mathbf{W}^O \quad (4.2)$$

avec

$$\begin{aligned} \text{head}_i &= \text{Attention}(\mathbf{Q}\mathbf{W}_i^Q, \mathbf{K}\mathbf{W}_i^K, \mathbf{V}\mathbf{W}_i^V) \\ \mathbf{W}_i^Q &\in \mathbb{R}^{d_{\text{model}} \times d_k}, \mathbf{W}_i^K \in \mathbb{R}^{d_{\text{model}} \times d_k}, \mathbf{W}_i^V \in \mathbb{R}^{d_{\text{model}} \times d_v} \end{aligned}$$

Les matrices \mathbf{W}_i^Q , \mathbf{W}_i^K , et \mathbf{W}_i^V sont des projections linéaires apprises qui transforment les séquences d'entrée en requêtes (\mathbf{Q}), clés (\mathbf{K}) et valeurs (\mathbf{V}) respectives pour chaque tête. Ces transformations permettent à chaque tête d'attention de se spécialiser dans des aspects spécifiques des relations contextuelles au sein des données. Enfin, les sorties de toutes les têtes ($\text{head}_1, \dots, \text{head}_h$) sont concaténées et projetées via \mathbf{W}^O pour produire l'attention finale.

Cette approche permet au modèle de capturer simultanément différentes relations complexes dans la séquence.

4.1.5.2 Architecture Transformer

L'architecture *Transformer* est composée de plusieurs blocs d'encodeurs et de décodeurs, chacun constitué de couches de *multi-head attention* et de couches *feed-forward* (propagation avant) (voir Figure 4.3).

L'encodeur (à gauche dans la Figure 4.3) traite une séquence d'entrée pour produire des représentations contextuelles à chaque position à l'aide des couches d'encodage avec une *multi-headed self-attention* (auto-attention multi-tête), tandis que le décodeur (à droite dans la

Figure 4.3) génère une nouvelle séquence en se basant sur ces représentations à l'aide des couches de décodeur avec une *masked multi-headed self-attention* (auto-attention multi-tête masquée) et une *multi-headed cross-attention* (attention croisée multi-tête).

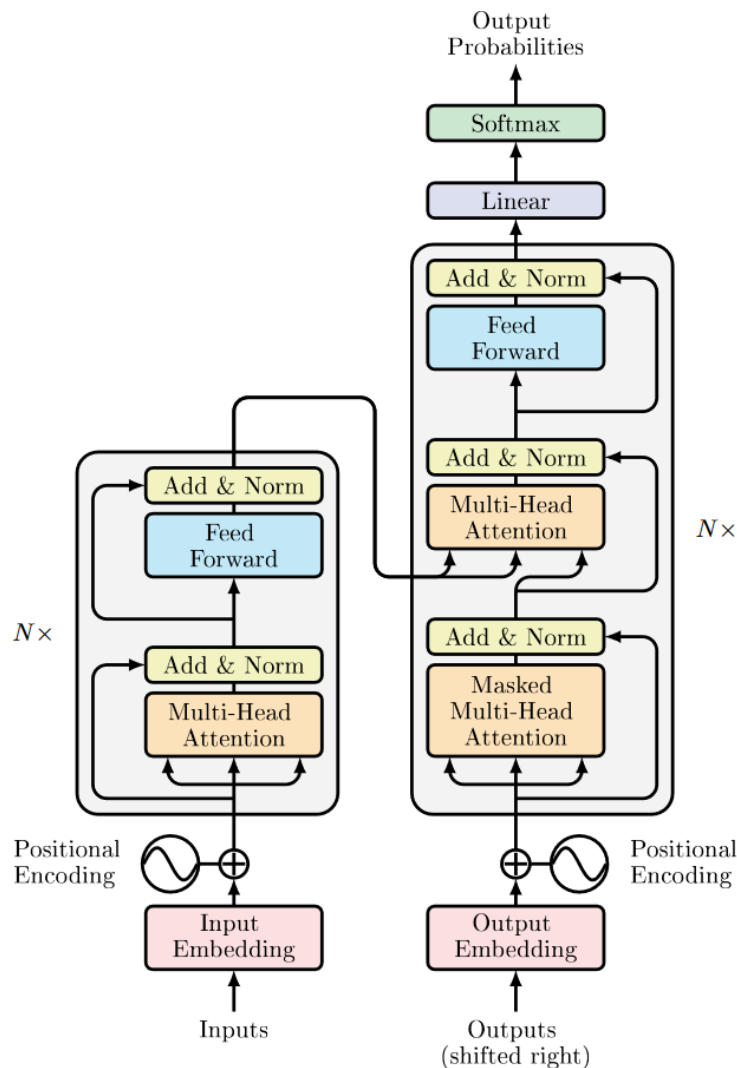


FIGURE 4.3 – Architecture du *Transformer*, avec un encodeur (à gauche) et un décodeur (à droite). N fait référence au nombre de couches [Vas+17].

Le décodeur intègre également un mécanisme de **cross-attention** (attention croisée) qui met en correspondance les séquences d'entrée et de sortie. Enfin, pour compenser l'invariance à la permutation des *Transformers*, un *positional encoding* (encodage positionnel) est ajouté à chaque *embedding* (incorporation, c.-à-d., plongement vectoriel) pour injecter des informations sur l'ordre des éléments :

$$PE(i, \delta) = \begin{cases} \sin\left(\frac{i}{10000^{2\delta/d}}\right) & \text{si } \delta = 2\delta' \\ \cos\left(\frac{i}{10000^{2\delta/d}}\right) & \text{si } \delta = 2\delta' + 1 \end{cases} \quad (4.3)$$

Cette architecture a démontré des performances supérieures dans de nombreuses tâches séquentielles, y compris la génération musicale. Elle a également été adaptée et optimisée dans plusieurs travaux ultérieurs pour améliorer l'efficacité et la qualité des résultats.

Ce modèle a fait ses preuves dans le domaine du *NLP* mais également dans de nombreux domaines. Cependant, comme tout modèle d'apprentissage profond, il est très sensible à la nature, la qualité, la diversité et la représentation des données d'entraînement. C'est pourquoi, dans nos travaux nous avons accordé toute notre attention au processus de traitement des données fournies en entrée, et plus particulièrement le processus de tokenisation de la musique symbolique (voir Section 4.1.3).

4.2 Positionnement de nos travaux de recherche

Nos travaux de recherche s'inscrivent d'une part dans une démarche de standardisation des techniques de tokenisation pour la musique symbolique et d'optimisation de ces mêmes techniques d'autres parts. Nos travaux visent à palier les limites identifiées dans ce domaine dans les approches existantes. Nos contributions ont donc porté sur l'unification des approches de tokenisation, sur la réduction par la compression de la longueur des séquences, et sur l'étude de l'impact de différentes tokenisation sur les performances des modèles d'apprentissage profond.

1. **Standardisation de la tokenisation** : *MidiTok*², créé spécifiquement pour faciliter la tokenisation des fichiers *MIDI*, propose une interface unifiée qui permet de comparer différentes méthodes de tokenisation, telles que *REMI* et *Compound Word*. *MidiTok* offre une grande flexibilité et permet aux chercheurs et ingénieurs d'entraîner et d'utiliser des modèles de langage pour la musique symbolique avec plus d'efficacité. Il répond à un besoin pressant de la communauté scientifique, en devenant la solution de référence pour les expérimentations sur la tokenisation de la musique symbolique. Nos travaux sur la standardisation de la tokenisation de musique symbolique [Fra+21] ont été publiés à la conférence ISMIR (International Society for Music Information Retrieval Conference) en 2021 (conférence de rang A).
2. **Réduction de la longueur des séquences** : Nous avons intégré la technique du *Byte Pair Encoding (BPE)* [BD20; Shi+99] pour la tokenisation de la musique symbolique, qui permet de compresser les séquences de tokens tout en préservant des éléments clés de la structure musicale. *BPE*, tout en réduisant efficacement la longueur des séquences, améliore la performance des modèles en termes de temps d'entraînement et d'inférence. Cependant, il convient de noter que cette réduction peut entraîner une

2. *Miditok* Repository : <https://github.com/Natooz/MidiTok>

perte de précision dans la représentation des événements musicaux individuels, notamment dans les nuances fines telles que la vitesse et la durée. Nos travaux sur le *BPE* [Fra+23a] ont été publiés à la conférence EMNLP (Empirical Methods in Natural Language Processing) en 2023 (conférence de rang A*).

- 3. Étude de l'impact des choix de tokenisation sur les performances des modèles :** Nos travaux ont montré que la manière dont les durées des notes et les événements temporels sont représentés influence directement les performances des modèles d'apprentissage profond. En examinant ces choix de tokenisation, nous avons démontré comment ils affectent non seulement la génération musicale, mais aussi d'autres tâches telles que la classification des compositeurs ou des émotions. Cela permet d'améliorer la flexibilité des modèles dans le traitement de structures musicales complexes (p.ex., la polyphonie). Nos travaux sur l'impact des choix de tokenisation [Fra+23b] ont été publiés à la conférence ISMIR (International Society for Music Information Retrieval Conference) en 2023 (conférence de rang A)

En conclusion, nos travaux apportent des contributions distinctes sur l'optimisation des représentations symboliques musicales pour l'apprentissage profond. Cela se traduit plus particulièrement par l'étude et l'amélioration de la tokenisation pour la musique symbolique : 1) *MidiTok* standardise et unifie les approches de tokenisation ; 2) *BPE* permet de réduire la complexité des séquences tout en optimisant les performances des modèles ; 3) Nos travaux sur l'impact des choix de tokenisation fournissent des informations précieuses pour améliorer les modèles d'apprentissage profond dans ce domaine.

4.3 Contribution logicielle : *Miditok*

Cette section vise à rappeler quelques travaux de l'état de l'art sur la tokenisation de la musique symbolique et à décrire notre contribution logicielle *Miditok*. Pour ce faire, elle repose sur des éléments présentés dans [Fra+21] et sur <https://github.com/Natooz/MidiTok>.

4.3.1 Avant *Miditok*

Les premiers travaux utilisant des modèles discrets pour la musique symbolique, tels que *DeepBach* [HPN17], *FolkRNN* [SSK15] ou *BachBot* [Lia+17], reposent sur des tokenisations spécifiques conçues pour les données utilisées, par exemple pour les quatre voix des chorals de Bach. Les modèles non-autorégressifs comme *MuseGAN* [Don+18] représentent souvent la musique sous forme de matrices de pianoroll. Depuis, les chercheurs ont introduit des tokenisations plus universelles adaptées à tout type de musique. Ces stratégies représentent les attributs des notes et le temps de manière générale, pouvant être utilisées avec n'importe quelle musique. Les plus couramment utilisées sont *Midi-Like* [Oor+18] et *REMI* [HY20]. La première tokenise la musique en représentant les événements du protocole *MIDI* sous forme

de tokens, tandis que la seconde représente le temps avec des tokens *Bar* et *Position* et la durée des notes avec des tokens explicites *Duration*. De plus, *REMI* inclut des tokens fournissant des informations supplémentaires comme les accords et le tempo.

Plus récemment, les chercheurs se sont concentrés sur l'amélioration de l'efficacité des modèles avec de nouvelles techniques de tokenisation p.ex., *Compound Word* [Hsi+21], *Octuple* [Zen+21] et *PopMAG* [Ren+20] fusionnent les vecteurs d'embeddings avant de les passer au modèle ; *LakhNES* [Don+19], [Pay19], *SymphonyNet* [Liu+22] utilisent des tokens combinant plusieurs valeurs, comme les hauteurs et les vitesses.

4.3.2 *Miditok* : paquet Python pour la tokenisation des fichiers *MIDI*

4.3.2.1 Une contribution logicielle

La tokenisation de la musique, c'est-à-dire la conversion des notes en séquences de tokens, n'est pas un processus simple. Contrairement au texte, la musique polyphonique comporte des notes simultanées, chacune avec plusieurs propriétés, et le problème devient encore plus complexe si l'on considère plusieurs pistes ou instruments. De nombreux articles de recherche récents ont introduit différentes manières de tokeniser la musique symbolique, mais peu d'auteurs partagent leur code source de manière à permettre une reproduction facile ou une simple utilisation de leurs méthodes. De plus, les fichiers musicaux, tels que les fichiers *MIDI*, doivent être correctement prétraités. Cette étape consiste à réduire la résolution de certaines valeurs, telles que les temps d'attaque et de relâchement des notes, leurs vitesses, leurs durées, ou encore les valeurs de tempo, entre autres.

Dans le but d'offrir un moyen simple et pratique de tokeniser les fichiers *MIDI*, nous avons créé un package python nommé *Miditok* [Fra+21]. Il implémente les tokenisations les plus populaires sous une API unifiée. Ces méthodes de tokenisations sont rappelées dans la Section 4.1.3.2 et comparées dans l'Annexe A.5. Les tokens supplémentaires en lien à ces méthodes sont rappelés dans l'Annexe A.6.

Miditok offre ainsi une grande flexibilité et des fonctionnalités étendues, permettant de former et d'utiliser facilement des modèles de langage pour la musique symbolique, et de comparer les différentes méthodes de tokenisation. Nous avons introduit *Miditok* pour la première fois fin 2021 [Fra+21], et il a depuis reçu plusieurs mises à jour jusqu'à devenir une référence dans la communauté. Aujourd'hui, *Miditok* est devenu une solution très utilisée par les chercheurs et ingénieurs (Logiciel téléchargé plus de 130,000 fois) souhaitant utiliser des modèles de langage avec la musique symbolique.

4.3.2.2 Tokeniser avec *Miditok*

Le workflow de tokenisation d'un fichier *MIDI* dans *Miditok* est le suivant :

1. Prétraitement de l'objet *MIDI* :
 - Si paramétré en mode "flux de tokens unique", fusion des pistes du même programme/instrument ;

- Suppression des notes avec des hauteurs en dehors de la plage du tokenizer ;
 - Rééchantillonnage de la division temporelle du *MIDI*, autrement dit rééchantillonnage des temps des messages représentant les notes, la signature rythmique, le tempo ou d'autres événements ;
 - Duplication des notes, des changements de tempo et de signature rythmique, et suppression des pistes vides ;
2. Création des tokens *MIDI* globaux (tempo, etc.) ;
 3. Création des tokens de piste (notes, accords) ;
 4. Si paramétré en mode "flux de tokens unique", concaténation de tous les tokens globaux et de piste, sinon concaténation des tokens globaux à chaque séquence de tokens de piste, puis tri par ordre d'occurrence ;
 5. Déduction des tokens temporels pour toutes les séquences de tokens (une seule si en mode "flux de tokens unique") ;
 6. Renvoi des tokens, sous forme d'une combinaison de liste de chaînes de caractères et de liste d'entiers (identifiants de tokens).

Notes : La première et la dernière étape sont réalisées de la même manière pour tous les tokenizers, tandis que les autres étapes peuvent être effectuées différemment selon la méthode de tokenisation. L'étape de prétraitement est essentielle car elle formate les informations d'un fichier *MIDI* pour qu'elles correspondent aux paramètres du tokenizer. Les temps de début et de fin des tokens de piste et globaux sont alignés sur la résolution temporelle du tokenizer, ainsi que leurs valeurs. Cela nous garantit de retrouver exactement le même fichier *MIDI* prétraité lors de la détokenisation des tokens.

4.4 Impact de la tokenisation du temps et de la durée des notes sur la modélisation de la musique symbolique

Nos travaux se sont concentrés sur l'analyse comparative des différentes stratégies de tokenisation du temps et de la durée des notes dans les modèles de musique symbolique. La motivation principale de cette étude était d'explorer les effets de ces choix de représentation sur diverses tâches de modélisation musicale.

En effet, les méthodes de tokenisation de la musique symbolique utilisées dans les travaux précédents sont souvent spécifiques à certains types de données musicales et peu d'études ont exploré l'impact comparatif de ces méthodes sur différentes tâches.

Les travaux que nous présentons dans cette section, comblent cette lacune en évaluant donc l'influence des représentations du temps et de la durée des notes. Cette section repose sur des éléments présentés dans [Fra+23b].

4.4.1 Méthodologie

4.4.1.1 Modèles et entraînements

Pour l'ensemble des expériences, nous avons utilisé l'architecture *Transformer* [Vas+17] avec des dimensions identiques : 12 couches, une dimension de 768 neurones, 12 têtes d'attention et des couches internes de 3072 neurones.

Pour les tâches de classification et de représentation séquentielle, le modèle est pré-entraîné sur 100 000 étapes avec un taux d'apprentissage de 10^{-4} , puis ajusté sur 50 000 étapes avec un taux d'apprentissage de 3×10^{-5} , et une taille de lot de 48 exemples.

Les modèles de génération reposent sur la bibliothèque *GPT-2* [Wol+20] et utilisent un masque d'attention causale (*Causal Attention Masking*) pour générer des séquences auto-regressivement.

4.4.1.2 Caractéristiques techniques

Les entraînements ont été effectués sur des *GPU V100* avec la précision mixte (AMP - Automatic Mixed Precision) [Mic+18] et l'optimiseur *Adam* [KB15].

Préparation des données d'expérimentation. Un pourcentage de 10% des données est utilisé pour la validation, et 15% pour tester les modèles. Les entrées contiennent entre 384 et 512 tokens, avec un token BOS pour le début de séquence et un token EOS pour la fin.

4.4.1.3 Tokenisations

Nous explorons quatre combinaisons possibles de représentation du temps et de la durée des notes à partir de *TS* (TimeShift), *Pos* (Position), *Dur* (Duration) et *NOff* (NoteOff) (Voir Section 4.1.2 pour leur définition.) :

1. *TS + Dur* ;
2. *TS + NOff* ;
3. *Pos + Dur* ;
4. *Pos + NOff*.

Les combinaisons *TS + Dur* et *TS + NOff* correspondent respectivement à *TSD* [Fra+23a] et *MIDI-Like* [Oor+18], tandis que *Pos + Dur* est équivalent à *REMI* sans tokens supplémentaires.

Nous appliquons différentes résolutions pour les tokens *Duration* et *TimeShift*, en échantillonnant plus précisément les notes courtes. Aucun token supplémentaire (par exemple, *Chord*, *Tempo*) n'est utilisé. Les données sont augmentées par des variations de hauteur et de vélocité, et la tokenisation est effectuée à l'aide de notre contribution logicielle *MidiTok* [Fra+21] (voir Section 4.3).

4.4.1.4 Tâches évaluées

Tâches de génération musicale. Pour la tâche générative, nous utilisons le jeu de données POP909 [Wan+20]. Les modèles commencent avec un prompt composé de 384 à 512 tokens, puis génèrent autoregressivement 512 tokens supplémentaires. L'évaluation des résultats générés reste une question ouverte [YL20]. Les travaux précédents mesurent souvent la similarité de certaines caractéristiques telles que la plage de hauteurs de note (pitch) entre les prompts et les résultats générés, en complément d'évaluations humaines. Cependant, cela n'est pas toujours très pertinent : un résultat généré peut avoir des caractéristiques très similaires à son prompt tout en étant de mauvaise qualité. Les évaluations humaines, bien qu'elles soient plus fiables pour juger de la qualité, peuvent aussi introduire des biais.

Ainsi, nous avons choisi d'évaluer les résultats à l'aide des ratios d'erreurs de prédiction : *Token Syntax Error* (TSE) [Fra+23a]. Cette métrique est exempte de biais et directement liée aux choix de conception des tokenisations. Elle nous permet de mesurer dans quelle mesure un modèle parvient à effectuer des prédictions fiables en se basant sur le contexte d'entrée et les connaissances qu'il a apprises.

Nous utilisons les catégories de [Fra+23a] :

- TSE_{type} : une erreur de type, par exemple lorsque le modèle prédit un token d'un type incompatible avec le précédent.
- TSE_{time} : une valeur `Position` mal prédite, qui revient en arrière ou reste dans le temps.
- TSE_{dupn} (note dupliquée) : une note prédite alors qu'elle était déjà en cours de lecture au temps actuel.
- TSE_{nof} (pas de NoteOff) : un token `NoteOn` prédit sans qu'un token `NoteOff` correspondant ne suive pour l'arrêter.
- TSE_{non} (pas de NoteOn) : un token `NoteOff` prédit alors que cette note n'était pas en cours de lecture.

Pour chaque token généré, une fonction basée sur des règles analyse son type et sa valeur afin de déterminer si ceux-ci sont valides ou, dans le cas contraire, de quel type d'erreur il s'agit. Le nombre total d'erreurs est normalisé par le nombre de tokens prédits.

Tâches de classification des compositeurs et des émotions. Pour certaines tâches de classification, la musique symbolique est sans doute mieux adaptée que l'audio ou les piano-rolls. Cela est particulièrement vrai pour la classification de caractéristiques de musique classique, comme l'identification du compositeur [KCW20]. La musique mono-instrumentale avec des mélodies et harmonies complexes, sans effets audio particuliers, bénéficie d'une représentation discrète pour les tâches de classification et de modélisation. Cela nous a semblé donc important pour mener des expériences sur ce type de tâche.

Nous avons ainsi choisi de porter nos expérimentations sur le jeu de données GiantMIDI [Kon+21] pour la classification de compositeurs et le jeu de données EMOPIA [Hun+21] pour la classification des émotions. La tâche de classification implique principalement de modéliser les motifs caractéristiques des compositeurs ou des émotions à partir des données.

Tâche de représentation des séquences. Cette tâche consiste à obtenir une représentation d'embedding de taille fixe pour une séquence d'entrée de tokens $p_\theta : \mathbb{V}^L \mapsto \mathbb{R}^d$. Ici, $\mathbb{V} \subset \mathbb{IN}$ désigne les identifiants des tokens du vocabulaire \mathcal{V} , L est la longueur variable de la séquence d'entrée, et d la taille des embeddings. En d'autres termes, le modèle apprend à projeter une séquence de tokens d'entrée dans un espace d'embedding, fournissant ainsi une représentation universelle. Nous trouvons cette tâche intéressante et bien adaptée pour évaluer les performances du modèle, car elle l'entraîne directement à modéliser les relations entre les tokens au sein de la séquence d'entrée et entre les différentes représentations elles-mêmes. Bien que les applications réelles de cette tâche pour la musique symbolique soient actuellement limitées, elle constitue une technique utile pour évaluer l'impact de la tokenisation sur l'apprentissage des modèles.

Cette tâche a déjà été abordée en traitement automatique du langage avec SentenceBERT [RG19] ou SimCSE [GYC21]. Nous avons adopté l'approche de ce dernier, qui utilise l'apprentissage contrastif pour entraîner le modèle à apprendre des représentations de séquences, où des entrées similaires ont des similarités de cosinus plus élevées. L'embedding de séquence est obtenu en effectuant une opération de pooling sur les états cachés de sortie du modèle. Nous avons décidé d'utiliser le dernier état caché de la position du token BOS, car cela a donné de bons résultats avec SimCSE [GYC21]³. Nous avons entraîné les modèles avec la méthode du dropout : pendant l'entraînement, un lot de n séquences $\mathcal{X} = \{\mathbf{x}_i\}_{i=0}^n$ est passé deux fois dans le modèle, mais avec des masques de dropout différents, ce qui donne des embeddings de séquences de sortie différents $\mathcal{Z} = \{\mathbf{z}_i\}_{i=0}^N$ et $\tilde{\mathcal{Z}} = \{\tilde{\mathbf{z}}_i\}_{i=0}^N$. Bien que le dropout modifie les sorties, la plupart des informations d'entrée restent accessibles au modèle. Par conséquent, nous nous attendons à ce que les paires d'embeddings de séquences $(\mathbf{z}_i, \tilde{\mathbf{z}}_i)$ soient similaires, avec une forte similarité de cosinus. Pour atteindre cet objectif, nous entraînons le modèle avec une fonction de perte définie par l'entropie croisée appliquée aux similarités de cosinus par paires au sein du batch (sim) :

$$\ell_i = -\log \frac{e^{\text{sim}(\mathbf{z}_i, \tilde{\mathbf{z}}_i) / \tau}}{\sum_{j=1}^N e^{\text{sim}(\mathbf{z}_i, \tilde{\mathbf{z}}_j) / \tau}} \quad (4.4)$$

Ainsi, le modèle apprend efficacement à créer des embeddings de séquences similaires pour des entrées similaires, tout en éloignant celles présentant des dissimilarités. Nous avons utilisé une valeur de dropout de 0,1 pour entraîner les modèles et utilisé le jeu de données GiantMIDI [Kon+21].

L'évaluation de la représentation de séquences est réalisée en mesurant les distances et similarités entre des paires de séquences similaires. Nous avons recours à l'augmentation des données en décalant la hauteur (pitch) et la vitesse des séquences afin d'obtenir des paires de séquences musicales similaires. Les données augmentées conservent la plupart des informations des données originales. Ainsi, les modèles produisent des embeddings similaires pour les paires de séquences originales-augmentées.

3. SimCSE utilise un token CLS, équivalent à BOS dans notre cas.

4.4.2 Synthèse des résultats

L'ensemble des expérimentations et résultats est disponible dans notre article [Fra+23b]. Ci-dessous nous en réalisons la synthèse. Les expérimentations que nous avons menées ont permis d'évaluer quatre combinaisons de représentations temporelles et de durée des notes sur plusieurs tâches :

1. **Tâches de génération musicale** : La combinaison `TimeShift` + `Duration` (TS + Dur) obtient les meilleurs résultats pour la génération musicale. (TS + Dur) réduit les erreurs de prédiction, évitant ainsi la duplication ou l'oubli des notes. Les tokens de `Position` tendent quant à eux à générer des erreurs temporelles, entraînant des chevauchements de notes.
2. **Tâche de classification des compositeurs et des émotions** : La combinaison `TimeShift` + `Duration` (TS + Dur) se démarque également dans les tâches de classification. La représentation plus explicite des décalages temporels et des durées permet de mieux capturer les motifs caractéristiques des compositeurs et des émotions.
3. **Tâche de représentation des séquences** : La tokenization basée sur `Position` montre un certain avantage dans les tâches d'apprentissage de représentation des séquences. C'est plus particulièrement le cas dans les tâches d'apprentissage contrastif où les similarités entre séquences doivent être identifiées.

Nous pouvons ainsi conclure que la représentation explicite de la durée des notes via des tokens `Duration` est plus performante pour la plupart des tâches, comparée à des méthodes implicites avec des tokens `NoteOff`. De plus, l'utilisation de tokens `TimeShift` pour représenter les mouvements temporels améliore la précision et réduit les erreurs, surtout dans les tâches de génération.

4.4.3 Discussion

Dans cette section réflexive sur la contribution des méthodes de tokenisation pour la modélisation de la musique symbolique en apprentissage profond, il est important de souligner l'impact des représentations explicites ou implicites du temps et des durées des notes.

Les différentes méthodes de tokenisation, telles que *REMI*, *MIDI-Like* ou encore des approches plus récentes comme *Compound Word* et *Octuple*, offrent chacune des avantages distincts selon les tâches. La principale réflexion que nous avons menée porte sur le choix de la meilleure approche de tokenisation en fonction des objectifs de modélisation. Par exemple, dans le cas de la génération musicale, les modèles peuvent bénéficier d'une meilleure représentation (plus explicite) du temps et des durées de note grâce à des tokens comme `TimeShift` et `Duration`, qui offrent une meilleure granularité dans la représentation des éléments temporels et rythmiques de la musique. Cela permet d'améliorer la précision et la qualité des résultats générés.

En revanche, pour d'autres tâches, comme la représentation des séquences, les représentations implicites basées sur des tokens `Position` peuvent se révéler pertinentes. Ces

représentations obligent le modèle à déduire les durées de notes à partir des combinaisons de tokens de début et de fin de note, ce qui semble bien adapté aux méthodes d'apprentissage contrastif, où la capacité à distinguer les nuances subtiles entre les séquences musicales est cruciale. Toutefois, les représentations explicites ont montré de bons résultats également dans ces tâches.

Notre recherche démontre qu'il n'existe pas une solution unique de tokenisation idéale pour toutes les tâches. En fonction des exigences de la tâche — génération, classification ou apprentissage contrastif — les méthodes de tokenisation doivent être soigneusement choisies. Par exemple, l'utilisation de tokens explicites est essentielle pour la génération musicale ou la classification de compositeurs ou la reconnaissance des émotions, permettant des résultats cohérents et musicaux. En revanche, la représentation implicite des durées de note, bien que demandant un effort supplémentaire au modèle, semble apporter une meilleure différenciation des séquences musicales.

De tels constats montrent l'importance de bien appréhender les caractéristiques spécifiques de chaque méthode de tokenisation et la manière dont elles influencent la performance des modèles en fonction des tâches spécifiques pour lesquelles on les emploie. L'un des facteurs clés pour améliorer la performance des modèles d'apprentissage de musique symbolique repose donc sur la capacité d'ajuster les méthodes de tokenisation en fonction des besoins, tout en tenant compte de la structure temporelle et de la nature des notes.

En conclusion, on pourra faire le lien notamment entre les travaux du Chapitre 1 et ce chapitre puisque de nouveau l'expression "There is no silver bullet" [BK87] prend tout son sens. Il devient alors primordial d'envisager des méthodes spécifiques et appropriées aux tâches que l'on souhaite réaliser. Dans notre étude on remarquera ainsi que la représentation du temps et de la durée des notes a un impact significatif sur les performances des modèles de musique symbolique. Les représentations explicites de la durée des notes sont à privilégier pour les tâches de classification et de génération, tandis que les positions des événements musicaux sont plus adaptées aux tâches de représentation des séquences. La combinaison `TimeShift` + `Duration` est recommandée pour la majorité des tâches, car elle offre un bon équilibre entre performance et fiabilité. Des recherches futures pourraient explorer d'autres dimensions de la tokenisation, comme la représentation multi-piste et l'intégration des métadonnées pour des tâches plus complexes, telles que la transcription musicale.

4.5 Byte Pair Encoding pour la musique symbolique

Cette section vise à décrire notre contribution sur l'incorporation du *Byte Pair Encoding* pour la musique symbolique. Pour ce faire, elle repose sur des éléments présentés dans [Fra+23a].

4.5.1 Contexte et motivation

Comme décrit dans les sections précédentes, la musique symbolique, utilisée dans les modèles d'apprentissage profond, doit être convertie en séquences de tokens discrets pour pou-

voir être traitée par des modèles de langage comme les *Transformers* [Vas+17]. Les approches classiques de tokenisation reposent sur des vocabulaires relativement limités, où chaque token représente des attributs individuels des notes comme la hauteur, la vitesse ou la durée. Cela entraîne des séquences de tokens longues, ce qui non seulement limite l'efficacité des modèles, mais sous-utilise également la capacité d'encodage des modèles de langage, souvent dimensionnés pour des vocabulaires beaucoup plus vastes.

4.5.2 Stratégies de réduction de la longueur des séquences

Des travaux récents se concentrent sur l'efficacité des modèles. Comme la plupart d'entre eux reposent sur l'architecture Transformer [Vas+17] et le mécanisme d'attention, leur complexité temporelle et spatiale augmente de manière quadratique avec la longueur de la séquence d'entrée. Ce goulot d'étranglement a conduit les chercheurs à travailler sur des estimations d'attention plus efficaces [Tay+21], jusqu'à atteindre une complexité linéaire. Dans le domaine de la musique symbolique en particulier, les chercheurs ont développé des stratégies visant à réduire la longueur des séquences afin d'augmenter 1) l'efficacité des modèles ; 2) l'étendue du mécanisme d'attention ; 3) la qualité des résultats générés. Ces stratégies peuvent être divisées en deux catégories :

- **regroupement d'embeddings** comme dans *Compound Word* [Hsi+21] (*CPWord*), *Octuple* [Zen+21], *PopMag* [Ren+20], *SymphonyNet* [Liu+22] ou *MMT* [Don+23]. Les embeddings de plusieurs tokens sont fusionnés par une opération de regroupement (pooling). Cela se fait souvent en concaténant les embeddings et en projetant le vecteur, ce qui donne un embedding agrégé de taille fixe.
- **combinaison de tokens** comme dans *MuseNet* [Pay19], *LakhNES* [Don+19] ou dans d'autres travaux récents [Liu+22 ; Thi+23], qui consiste à utiliser un vocabulaire de tokens représentant plusieurs attributs, par exemple, *Pitch-x_Dur-y* représentant à la fois l'information de hauteur (pitch) et de durée (duration).

Dans le cadre de nos travaux, nous avons décidé d'adapter le *Byte Pair Encoding (BPE)* à la tokenisation de la musique symbolique afin de palier les problèmes des séquences trop longues. Aucune contribution positionnant le *BPE* dans le domaine de la musique n'avait abouti jusque-là. En revanche, nous considérons que le *BPE* peut être interprété comme une technique de combinaison de tokens apprise.

4.5.3 *Byte Pair Encoding (BPE)*

Le *Byte Pair Encoding (BPE)* [Gag94] est une technique de compression de données initialement utilisée dans le traitement automatique des langues pour segmenter les mots rares en sous-unités plus fréquentes. Elle convertit les successions de *bytes*⁴ les plus récurrentes dans un corpus en de nouveaux *bytes* créés. Par exemple, dans la séquence de caractères *aabaabaacaa*, la sous-séquence *aa* apparaît trois fois et est la plus récurrente. Apprendre et

4. caractères distincts qui composent le texte

appliquer le *BPE* sur cette séquence remplacerait *aa* par un nouveau symbole, par exemple *d*, donnant la séquence compressée *dbdbdcd*. Cette dernière peut être encore réduite en remplaçant la sous-séquence *db*, donnant *eedcd*. En pratique, le *BPE* est appris sur un corpus jusqu'à ce que le vocabulaire atteigne une taille cible. L'apprentissage du *BPE* est décrit par le pseudo-code de l'algorithme 7.

Algorithme 7 : Algorithme d'apprentissage du *BPE*

Entrée : Vocabulaire de base \mathcal{V} , taille cible du vocabulaire N , dataset \mathcal{X}
Sortie : Le vocabulaire appris \mathcal{V}

Tant que $|\mathcal{V}| < N$

1. Trouver $m = \{t_1, t_2\}$, la succession de tokens la plus récurrente dans \mathcal{X}
2. Mettre à jour le vocabulaire : $\mathcal{V} \leftarrow \mathcal{V} + [t_{|\mathcal{V}|} : m]$
3. Remplacer les occurrences de m dans \mathcal{X} par $t_{|\mathcal{V}|}$

Fin tant que

Retourner le vocabulaire appris \mathcal{V}

Le *BPE* est aujourd'hui largement utilisé dans le domaine du traitement du langage naturel (*TALN*, *Natural Language Processing - NLP*) car il permet d'encoder des mots rares et de segmenter des mots inconnus ou composés en séquences d'unités de « sous-mots » [SHB16]. D'autres techniques d'agrégation de tokens ou de construction de vocabulaire existent. Les deux autres techniques les plus couramment utilisées sont *Unigram* [Kud18] et *WordPiece* [Wu+16], dont les opérations partagent des similarités avec le *BPE*.

Pour le langage naturel, les *bytes* sont les caractères distincts qui composent le texte. Pour la musique symbolique, les attributs distincts de notes et de temps peuvent être utilisés comme des *bytes* à fusionner. Dans ce contexte, le *BPE* permet de représenter une note, ou même une succession de notes, récurrente dans le jeu de données, sous forme d'un seul token. Par exemple, une note qui serait tokenisée comme une succession de tokens *Pitch_D3*, *Velocity_60*, *Duration_2.0* pourrait être remplacée par un seul token nouveau. Les notes rares (et leurs attributs) peuvent toujours être tokenisées sous forme de tokens non-*BPE*. La même logique s'applique aux tokens temporels, qui peuvent également être associés aux tokens de notes.

Nous avons donc proposé l'application du *BPE* à la musique symbolique comme une solution permettant de compresser les séquences de tokens tout en augmentant la taille du vocabulaire et en rendant les représentations plus expressives. Le *BPE* apprend à combiner les séquences récurrentes de tokens en nouveaux tokens, ce qui permet de réduire considérablement la longueur des séquences sans perdre d'information musicale.

4.5.4 Synthèse des résultats

Les expérimentations menées sur plusieurs tâches de génération et de classification montrent que l'application du *BPE* à la musique symbolique présente plusieurs avantages :

- La réduction significative de la longueur des séquences de tokens, pouvant aller jusqu'à plus de 50%. Ceci améliore l'efficacité des modèles en termes de besoin en *GPU*.
- L'amélioration des performances des modèles, aussi bien pour la génération de musique que pour la classification, et ce grâce à l'utilisation d'un vocabulaire plus adapté voire plus expressif.
- L'accélération des temps d'inférence. Nous avons observé des gains de vitesse jusqu'à deux fois inférieurs aux méthodes sans *BPE*.

Enfin, l'un des effets positifs majeurs de l'utilisation de *BPE* réside dans l'amélioration de la distribution des embeddings dans l'espace de représentation. En augmentant la taille du vocabulaire tout en réduisant la longueur des séquences, *BPE* permet aux modèles d'apprendre des représentations plus isotropes et mieux réparties dans l'espace d'embedding. Cela optimise l'utilisation des capacités d'apprentissage du modèle et conduit à des performances accrues dans les tâches de génération et de classification.

4.5.5 Synthèse et discussion

Notre contribution propose l'utilisation et l'incorporation du *Byte Pair Encoding (BPE)* dans la tokenisation de la musique symbolique. Cette méthode ayant été utilisée initialement en *NLP*, le *BPE* est ici appliqué pour réduire la longueur des séquences musicales tout en conservant la richesse de l'information musicale. Les principales contributions que nous avons proposées sont les suivantes :

- le *BPE* a permis de fusionner plusieurs tokens en un seul. Ceci permet d'améliorer l'efficacité des modèles (basés sur les *Transformers*) ;
- le *BPE* s'est adapté à différentes représentations musicales, y compris les notes, les durées et les tempos ;
- les modèles utilisant le *BPE* ont été plus efficaces et précis, avec une meilleure gestion des longues séquences musicales ;
- le *BPE* se distingue de manière positive par rapport à d'autres approches telles que *MIDI-Like* ou *REMI*.

L'approche *BPE* est une contribution dans le domaine de la modélisation de la musique symbolique en ce qui concerne la gestion de la longueur des séquences. En adaptant cette technique issue du *NLP* au domaine musical, nous avons démontré la polyvalence de certaines méthodes de traitement des séquences. Le *BPE* a non seulement amélioré l'efficacité des modèles, mais a également permis de traiter des morceaux plus longs et plus complexes.

Cependant, des problématiques subsistent, notamment pour la gestion de la polyphonie et des événements musicaux complexes. Bien que le *BPE* soit une solution prometteuse, il est essentiel de s'assurer que la fusion des tokens ne conduit pas à la perte d'informations musicales cruciales.

Ces perspectives de recherche mettent en avant l'importance de trouver un équilibre entre réduction de séquence et préservation des informations musicales essentielles. Ainsi, dans le cadre de futures recherches, il serait intéressant de se concentrer sur l'optimisation de ces algorithmes pour mieux capturer les nuances de la musique comme l'expressivité et la dynamique.

En conclusion, le *BPE* constitue une avancée pour la tokenisation de la musique symbolique. En offrant un compromis efficace entre la taille du vocabulaire et la longueur des séquences, cette technique permet d'améliorer à la fois les performances des modèles et l'efficacité des temps d'inférence. De futures recherches pourraient explorer d'autres techniques de construction de vocabulaires, comme Unigram [Kud18] ou WordPiece [Wu+16], pour déterminer s'il est possible d'optimiser encore davantage la tokenisation pour la musique symbolique.

4.6 Réflexion critique générale sur les LLMs

Dans ce chapitre, nous avons porté des modèles issus du *NLP* pour effectuer entre autres des tâches de génération de musique symbolique. Or, nous restons convaincus que ces modèles restent pourtant imparfaits. Ci-dessous, nous en dressons quelques limites et évoquons de nouvelles pistes de recherche.

Les limites du raisonnement des modèles de langage L'évolution récente des modèles de langage, notamment dans leur capacité à résoudre des problèmes complexes en apparence, a suscité beaucoup d'enthousiasme sur leur potentiel à « raisonner ». Cependant, à l'examen plus attentif, il apparaît que ces modèles ne font pas preuve d'un véritable raisonnement formel [Mir+24]. Ils ne possèdent pas une compréhension conceptuelle des tâches, mais s'appuient sur une forme de correspondance probabiliste de motifs (pattern matching), comme l'ont démontré plusieurs travaux récents, y compris [Mir+24].

L'article *GSM-Symbolic : Understanding the Limitations of Mathematical Reasoning in Large Language Models* [Mir+24] met en lumière la fragilité des modèles face à des variations dans les problèmes posés, notamment en mathématiques. L'étude montre que de petits changements dans les paramètres d'un problème — même lorsqu'ils sont insignifiants du point de vue du raisonnement humain — entraînent des variations significatives dans les performances des LLMs. Ceci illustre que le modèle ne résout pas les problèmes en raisonnant sur les concepts sous-jacents, mais en reproduisant des motifs vus dans les données d'entraînement [Mir+24]. Cette observation s'accorde avec les constats de Yann Lecun (Prix Turing 2019) sur la manière dont les LLMs traitent l'information.

L'illusion du raisonnement et l'effondrement de l'entropie Comme nous en avons fait le constat dans certains de nos travaux courant (non publiés) et de nos expérimentations, l'effondrement de l'entropie dans les réponses des modèles suggère une forme de structure ordonnée qui n'est pas le fruit d'un raisonnement logique. Cet effondrement reflète une réduction de

la diversité des solutions générées par le modèle, ce qui pourrait indiquer que ces systèmes, au lieu de réellement réfléchir aux différentes possibilités, s'alignent sur les réponses les plus probables qu'ils ont rencontrées. On constate de ce fait une perte de diversité créative lorsque le modèle converge vers un certain type de réponse. Nos expériences de création d'un premier *Large Music Model* (non publiées et non évoquées dans ce chapitre) a aboutit notamment à un sur-apprentissage très important.

Les résultats obtenus dans l'étude [Mir+24] montrent que les *LLMs* sont extrêmement sensibles aux changements dans les clauses d'un problème. De telles considérations suggèrent qu'ils ne possèdent pas une logique interne robuste. Dès que l'on introduit des clauses additionnelles ou des chiffres différents, leurs performances chutent jusqu'à 65% dans certains cas. Ces modèles restent donc incapables de gérer l'augmentation de la complexité exactement de la même manière que les modèles autoregressifs ne peuvent raisonner que dans un cadre très restreint, imposé par leurs données d'entraînement.

Les implications pour la recherche future À la lumière de ces observations, il devient de plus en plus évident que les modèles de langage actuels ne raisonnent pas au sens traditionnel du terme. Ils sont doués pour la recherche et la correspondance de motifs, mais ils échouent dès que des éléments non pertinents ou perturbateurs sont introduits dans le problème, comme l'a illustré l'introduction du benchmark *GSM-NoOp* [Mir+24]. Leurs erreurs, comme interpréter des informations non pertinentes (par exemple, des valeurs de réduction) en des opérations mathématiques à part entière, montrent une incompréhension fondamentale des concepts qu'ils manipulent.

L'ensemble de ces constats impose de revoir notre manière d'évaluer le raisonnement des *LLMs*. Ils soulignent aussi de la nécessité de développer des modèles qui possèdent une capacité de raisonnement formel plus robuste. Il devient donc impératif d'élargir les recherches dans cette direction afin de palier les problématiques posés par la nature probabiliste (ou statistique) et non logique des *LLMs* actuels.

TROISIÈME PARTIE

Bilan et perspectives

CONCLUSION ET PERSPECTIVES DE RECHERCHE

5.1 Bilan de mes recherches

Ce mémoire présente une synthèse des contributions apportées au domaine de l'apprentissage automatique pour la génération et l'exploration de données complexes. Ces travaux se sont particulièrement concentrés sur les problématiques d'adaptation et de contextualisation. Je résumé ici les principaux apports des quatre chapitres.

Apprentissage par renforcement et systèmes de recommandation multicritères

Dans le **Chapitre 1**, nous avons abordé les problématiques d'optimisation multicritère dans les systèmes de recommandation au travers de l'utilisation de l'apprentissage par renforcement. Nous avons introduit de nouvelles méthodes utilisant les bandits-manchots contextuels et combinatoires pour améliorer à la fois la précision et la diversité des recommandations dans des environnements non stationnaires et avec des données parcimonieuses. L'une de nos contributions clé correspond à la suite de méthodes *Gorthaur* que nous avons développée. Les méthodes *Gorthaur* améliorent les performances des systèmes de recommandation bi-objectifs en environnement non stationnaire et sans connaissances a priori sur la nature des données. Les approches *Gorthaur* équilibrent efficacement l'exploration continue de nouvelles stratégies et l'exploitation des algorithmes optimaux. Elles présentent également des garanties théoriques, renforçant de ce fait la robustesse de ces méthodes dans le contexte de la recommandation.

Optimisation multicritère pour la sélection d'attributs médicaux

Dans le **Chapitre 2**, nous nous sommes concentrés sur l'optimisation multicritère appliquée à la sélection d'attributs dans le domaine médicale. Pour ce faire, nous avons développé un algorithme génétique, combiné à une méthode de calcul de fitness multicritère, pour résoudre des problèmes de sélection d'attributs, améliorant ainsi les processus décisionnels en médecine c.-à-d., maximisation de l'*AUC*, maximisation de la précision et minimisation du nombre d'attributs nécessaires, impliqués dans ce processus. Ce travail s'inscrit dans la continuité de l'optimisation multicritère présentée dans le **Chapitre 1**, en ciblant spécifiquement la réduc-

tion de la dimensionnalité tout en maintenant les performances prédictives des modèles et leur explicabilité.

Évaluation de comportements dynamiques non linéaires appliquée aux véhicules terrestres

Dans le **Chapitre 3**, nous nous sommes focalisés sur l'application de l'apprentissage profond à l'estimation des paramètres dans les systèmes dynamiques non linéaires, notamment ceux des véhicules terrestres. Nous avons proposé une méthode hybride combinant l'apprentissage profond (Perceptron Multicouche - *MLP*) et l'inférence bayésienne pour améliorer la précision des estimations de la dynamique latérale des véhicules. Cette approche a montré une grande pertinence dans la modélisation des systèmes complexes en intégrant des informations physiques, permettant ainsi une meilleure estimation et prédiction des comportements dynamiques. De plus, nous avons introduit une seconde contribution visant à améliorer l'estimation d'état en temps réel. Cette méthode repose sur l'utilisation de l'apprentissage profond (*GRU*, *CNN*, *CNN-T*) dans le cadre de l'estimation à horizon glissant (*MHE* - *Moving Horizon Estimation*), permettant une meilleure précision dans la résolution des problèmes d'optimisation en temps réel. Le modèle est capable de fournir des estimations rapides et fiables, même en présence de divergences importantes entre les modèles et les données réelles. Cette contribution peut également ouvrir des perspectives dans des applications en robotique ou dans les véhicules autonomes.

Optimisation des représentations symboliques musicales pour l'apprentissage profond

Enfin, dans le **Chapitre 4**, nous avons exploré l'utilisation de modèles d'apprentissage profond à base de *Transformers* pour la génération et la classification de musique symbolique. Nos travaux ont mis en avant l'importance du processus de tokenisation des séquences musicales pour améliorer la qualité des compositions générées. Notre contribution a été de proposer d'intégrer le *Byte Pair Encoding (BPE)* comme stratégie de réduction de la longueur des séquences, afin d'optimiser les performances des modèles. Ces travaux ont permis d'améliorer l'efficacité des modèles génératifs dans le domaine de la création musicale automatique, en offrant une meilleure gestion des séquences longues et une représentation plus compacte des données musicales. En parallèle, nous avons développé une contribution logicielle : **MidiTok**, une bibliothèque open-source dédiée à la tokenisation des fichiers *MIDI*. Cet outil a permis de standardiser et d'optimiser les différentes méthodes de tokenisation pour la musique symbolique, facilitant ainsi les expérimentations et la comparaison des approches du domaine.

Conclusion

Ces quatre chapitres illustrent la diversité des contributions apportées, allant de l'optimisation multicritère appliquée aux systèmes de recommandation et à la sélection d'attributs

médicaux, jusqu'à l'apprentissage profond pour l'estimation de paramètres dans les systèmes dynamiques non linéaires ou encore la génération musicale. Ces travaux mettent en lumière ma capacité à aborder des problématiques complexes liées :

- aux données en elles-mêmes par rapport à des caractéristiques variées telles que leur parcimonie, leur dimensionnalité (grande dimension), ou encore leur représentation ;
- aux environnements dynamiques ou non stationnaires.

En développant des méthodes d'apprentissage automatique ou de représentation de données d'entraînement pour la génération ou l'exploration, les recherches présentées dans cette *HDR* démontrent des approches adaptées, adaptables et robustes pour résoudre des problèmes complexes dans des domaines d'application variés. Cette démarche met en évidence une maîtrise des techniques d'apprentissage, aussi bien pour la modélisation complexe que pour l'optimisation des processus décisionnels et créatifs dans divers contextes.

5.2 Perspectives de recherche

Mes travaux de recherche de la Partie I, principalement axés sur l'apprentissage par renforcement et l'apprentissage automatique, ont montré une capacité à résoudre des problèmes multicritères. Ceux de la Partie II se sont focalisés sur l'apprentissage profond pour l'estimation de paramètres et pour des tâches de génération de musique symbolique. En s'appuyant sur des techniques d'apprentissage par renforcement, d'optimisation et d'apprentissage profond, j'ai ainsi abordé des thématiques allant des systèmes de recommandation, des systèmes dynamiques non linéaires à l'Intelligence Artificielle générative.

Les résultats obtenus jusqu'à présent offrent une base solide pour explorer de nouvelles problématiques de recherche notamment dans le cadre de mon projet récent : G -*GENOCOD*¹.

De plus, la thèse de Fabien Lionti avec l'INRIA et la DGA est en cours de finalisation et donne des perspectives de recherche dans le domaine l'Intelligence Artificielle informée/guidée par la physique, entre autres les capteurs virtuels ou encore la modélisation des roulis. En extension, il serait aussi possible de monter d'autres sujets sur cette thématique en dehors des applications spécifiques à l'armée ou à la dynamique des véhicules terrestres.

Cette section présente les travaux en cours et perspectives de recherche que je compte poursuivre dans les années à venir.

5.2.1 Travaux en cours et perspectives liées au projet G -*GENOCOD*

5.2.1.1 Contexte applicatif et genèse du projet

Contexte du projet. Dans divers secteurs de la chimie, la découverte de nouvelles molécules provient souvent de l'optimisation de composés existants, améliorant leurs propriétés par le biais de réactions chimiques. La création d'une molécule entièrement inédite est un événement plus rare, mais peut conduire à des optimisations et des transformations significatives dans le

1. Projet financé par la région Pays de la Loire dans le cadre du dispositif Étoiles Montantes.

domaine. C'est précisément cet objectif qui a orienté une recherche axée sur la génération *de novo* de molécules possédant des propriétés spécifiques, notamment dans le contexte de la découverte de médicaments et de matériaux [Leg+22]. Parmi les principaux enjeux de cette recherche figurent l'immensité de l'espace à explorer et la difficulté de générer des molécules susceptibles d'être synthétisées [CLD23].

Historique et genèse du projet. Les molécules peuvent être représentées sous forme de graphes, où les sommets symbolisent les atomes et les arêtes les types de liaisons. Cela pose un problème de génération de structures graphiques, où l'objectif est d'optimiser une ou plusieurs fonctions tout en respectant des contraintes spécifiques. Pour répondre à ces problématiques en chimie, le LERIA et Moltech-Anjou ont développé un algorithme évolutionnaire nommé EvoMol [Leg+20]. Ce générateur explore efficacement l'espace chimique et s'adapte à divers problèmes. Il a obtenu des résultats de référence sur des benchmarks d'optimisation multi-propriétés et des cas d'application concrets. De plus, il est capable d'intégrer des contraintes de synthétisabilité [CLD23] tout en favorisant la diversité des molécules générées [Leg+21].

Cependant, la nécessité d'intégrer un contexte moléculaire pour enrichir le raisonnement sur des variables de décision plus complexes, ainsi que l'amélioration de la génération de molécules « synthétisables », nous a conduits à proposer un nouveau projet : G -**GENOCOD**, ou **Graph-Generation for NOvel Compound Discovery**.

5.2.1.2 Objectifs du projet G -GENOCOD

Bien que le projet G -GENOCOD soit appliqué au domaine de la chimie, il aborde un problème plus général de génération de structures graphiques complexes par apprentissage par renforcement. Le cas spécifique en chimie correspond à un problème de génération de graphes dans un espace d'actions très vaste et composé. Dans la communauté de l'apprentissage par renforcement (*RL : Reinforcement Learning*) [LM96], ce type de problème soulève plusieurs problématiques à surmonter et pourrait également être utile dans d'autres domaines. Voici quelques-uns des principaux verrous que nous avons identifiés :

- le contexte moléculaire qui, par nature, pose deux problématiques majeures : 1) sa très haute dimensionnalité ; 2) la perte d'informations et potentiellement de sens pour l'apprentissage lors des phases de transformation (sous forme de fingerprints) et de préparation nécessaires pour le fournir sous forme de représentation structurée à un modèle d'apprentissage ;
- le caractère hautement combinatoire de la création d'une molécule, qui nécessite une séquence d'actions complexes ;
- le fait qu'une séquence d'actions ne puisse être récompensée qu'une fois la séquence complète réalisée, c'est-à-dire lorsque toutes les actions ont été effectuées et que les contextes intermédiaires ont été capturés ;
- les récompenses parcimonieuses (*sparse rewards*), qui rendent l'apprentissage difficile.

De plus, en chimie théorique, la fonction objectif est souvent une boîte noire dont l'évaluation est longue. L'un des enjeux est donc de développer des méthodes de *RL* capables de passer d'un problème simple à évaluer à un problème plus complexe, tout en améliorant le temps de calcul et la faisabilité. En effet, le *RL* nécessite souvent un grand nombre de données et d'itérations pour converger efficacement, ce qui est particulièrement contraignant dans le contexte chimique en raison de la haute dimensionnalité de l'environnement. C'est pourquoi nous travaillons sur l'utilisation de techniques comme l'apprentissage par transfert (*Transfer Learning*) [WKW16]. Mais, tous ces modèles, pré-entraînés puis affinés, doivent reposer sur des descripteurs utiles et pertinents du contexte moléculaire. Nous avons pu observer ces limites dans le cadre de nos travaux préliminaires

5.2.1.3 Correspondances avec les approches présentées dans les chapitres précédents

Comme mentionné dans la partie présentation synthétique de ce mémoire, le projet ^G-*GENOCOD* exige la mise en œuvre de plusieurs domaines et approches sur lesquels j'ai contribué au cours des cinq dernières années (Voir les quatre précédents chapitres). Parmi ces contributions, plusieurs se révèlent particulièrement pertinentes pour ce projet, notamment dans les axes suivants :

- l'utilisation d'algorithmes d'apprentissage par renforcement (liens avec le Chapitre 1) employés dans ce projet pour sélectionner les actions de mutations optimales à appliquer dans un graphe moléculaire durant la génération. De manière plus ambitieuse nous cherchons également par extension à déterminer des suites d'actions ordonnées à appliquer sur les molécules (voir Chapitre 1 sur les bandits-manchots combinatoires) ;
- l'utilisation d'un algorithme évolutionnaire pour la génération du graphe moléculaire et la prise en considération d'une nécessaire optimisation potentiellement multi-objectif (liens avec le Chapitre 1 et avec le Chapitre 2) ;
- le traitement et le raisonnement sur des données complexes (liens avec le Chapitre 4 sur l'optimisation des représentations et avec le Chapitre 2 sur la sélection d'attribut multicritère) qui seront utiles dans ce projet pour traiter et raisonner sur des descripteurs moléculaires complexes (p.ex., une meilleure manière de représenter les données de contexte moléculaire que les Extended-Connectivity Fingerprints (ECFPs) [RH10] ou une innovation en sélection de descripteurs moléculaires mono/multi-objectif) ;
- la potentielle prise en considération d'un modèle physique sous-jacent pour l'apprentissage (liens avec le Chapitre 3) ;
- l'exploration d'autres types d'approches d'IA générative (liens avec le Chapitre 4), également mises en œuvre dans ce projet pour la synthèse moléculaire. Ceci inclut notamment des méthodes d'apprentissage profond orientées vers la génération (liens avec le Chapitre 4), et plus spécifiquement les modèles de type *Transformer*.

5.2.2 Genèse des *Work Packages* de ^G-GENOCOD

Travaux préliminaires. Les travaux préliminaires que nous avons réalisés ont montré l'intérêt d'utiliser des algorithmes de *RL*. Nous avons tout d'abord expérimenté leur pertinence en leur apprenant à générer des molécules réalistes tout en optimisant le score de synthétisabilité (SAscore : Synthetic Accessibility Score) [ES09]. Cependant, ces premiers résultats ont été limités par un manque de contexte moléculaire (c'est-à-dire des caractéristiques chimiques ou variables contextuelles de la molécule) fourni par *EvoMol*. En effet, cette première version du générateur de molécules n'était pas conçue pour fournir un contexte moléculaire suffisamment riche sur lequel les méthodes d'apprentissage automatique pourraient s'appuyer pour résoudre le problème de la sélection des actions optimales sur le graphe (par exemple, les bandits-manchots contextuels, le *Q-Learning* et autres méthodes).

C'est pourquoi, nous nous sommes ensuite intéressés à l'acquisition d'un contexte moléculaire.

Le contexte moléculaire : *Work Package* - WP 1. Nos premiers travaux ayant montré qu'il était important de reposer sur un contexte moléculaire, nous avons décidé dans un premier temps de reposer sur les *Extended-Connectivity Fingerprints (ECFPs)* [RH10]. L'*ECFP* est un type de descripteur moléculaire couramment utilisé pour représenter des structures moléculaires de manière numérique dans des applications telles que la recherche de similarité, l'optimisation de molécules, ou encore la découverte de médicaments. Ces descripteurs sont utilisés principalement dans le domaine de la chemo-informatique pour encoder les caractéristiques structurales des molécules.

L'étude des descripteurs moléculaires fait aujourd'hui l'objet d'un *Work Package (WP)* du projet ^G-GENOCOD : le **WP 1**.

Réalisme des molécules : *Work Package* - WP 2 (Première étape). Des travaux supplémentaires ont ensuite été conduits dans un projet de Master afin d'apprendre à un algorithme de *RL (Q-Learning)* de choisir les actions optimales permettant de générer des molécules réalistes en reposant sur les *ECFPs*. Cet algorithme de *RL* choisit ainsi parmi trois actions de mutation : « Ajouter atome », « Retirer atome », « Modifier une liaison ».

Les premiers résultats que nous avons obtenus sont encourageants en ce qui concerne la capacité du *RL* à choisir les actions optimales permettant de générer des molécules réalistes (voir Figure 5.1). La Figure 5.1 représente le nombre d'échecs à générer des molécules réalistes selon deux méthodes de sélection des actions de mutation dans le graphe moléculaire : la méthode d'origine d'*EvoMol* (Mutations aléatoires) et la méthode employant de l'apprentissage par renforcement (*Q-Learning*) pour choisir l'action optimale de mutation du graphe moléculaire.

Les résultats étant prometteurs, nous poursuivons dans cette voie au travers d'un *Work Package* dédié du projet ^G-GENOCOD : le **WP 2**.

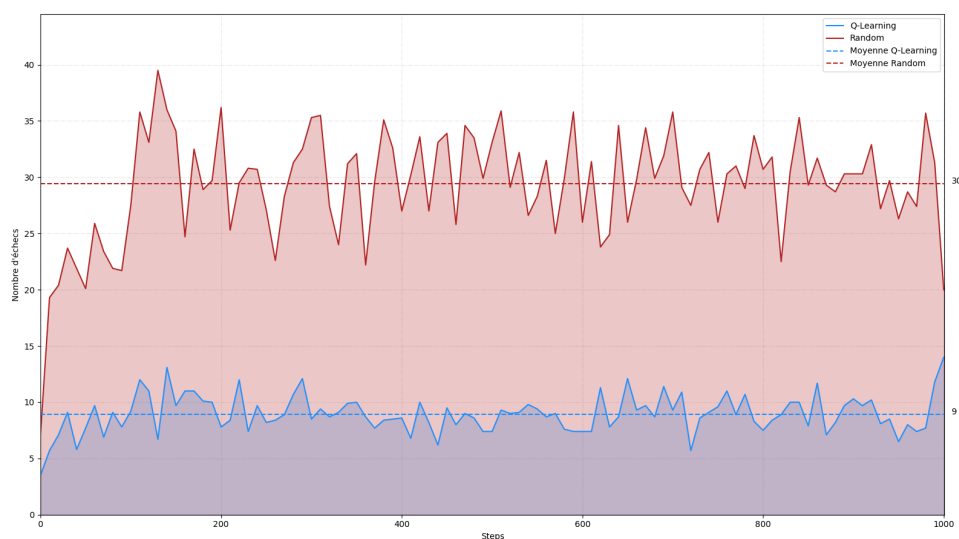


FIGURE 5.1 – Nombre d'échecs de génération de molécules réalistes sur 1000 itérations avec l'utilisation d'EvoMol d'origine (Random) versus l'utilisation du Q-Learning (reposant sur des *ECFP-0*) dans EvoMol

En revanche, même si les résultats obtenus avec les *ECFPs* sont encourageants, ces descripteurs sont loin d'être parfaits. Il est important de noter que l'une des problématiques actuelles est que ce contexte moléculaire est difficile à capturer, car difficile à décrire ou trop vaste. De plus, l'usage des *ECFPs* provoque une perte d'informations, d'explicabilité ou de sens. Nous continuons donc de réfléchir en parallèle du *WP 2*, aux descripteurs contextuelles des molécules (*WP 1*) à fournir aux méthodes d'apprentissage.

Transfer Learning : Work Package - WP 2 (Seconde étape). L'objectif final clé du *WP 2* restera néanmoins à ce que les algorithmes de *RL* que nous développons soient affinés (*fine-tuning*) ensuite pour aborder un problème d'optimisation différent du réalisme des molécules, tel que p.ex., l'optimisation de la QED (Quantitative Estimate of Drug-likeness) [Bic+12]. C'est à ce titre que nous envisageons l'utilisation de l'apprentissage par transfert une fois que les modèles de *RL* auront appris à choisir les actions permettant de générer des molécules réalistes.

Synthétisabilité : Work Package - WP 3. Enfin, nous avons identifié un troisième *Work Package (WP 3)* du projet *G-GENOCOD*. Il se concentre sur la synthétisabilité moléculaire à l'aide d'une approche *bottom-up*, qui permet de transformer une liste de réactions chimiques connues en actions applicables à des graphes moléculaires. Contrairement aux approches traditionnelles *top-down* de rétrosynthèse, complexes et souvent peu généralisables, cette approche permet d'apprendre à séquencer les réactions pour générer des molécules synthétisables et optimiser des propriétés cibles. Aucun travail préliminaire n'a été effectué pour proposer ce *WP*

3, en revanche, nous avons affecté Derin Özer sur ce sujet dans le cadre de son doctorat en informatique courant jusqu'à fin 2026.

5.2.2.1 Synthèse des Work Packages en cours du projet ^G-GENOCOD et perspectives

Le projet ^G-GENOCOD ouvre plusieurs perspectives pour l'optimisation de graphes moléculaires et la découverte de nouvelles molécules synthétisables.

Work Package 1 : Contexte Moléculaire. Le **Work Package 1** se concentre sur l'acquisition du contexte moléculaire. Cela inclut la capture de variables contextuelles pertinentes qui sont indispensables pour guider les algorithmes d'apprentissage par renforcement dans le choix des actions appliquées sur les graphes moléculaires. Ces variables englobent des descripteurs moléculaires globaux et des caractéristiques locales spécifiques aux actions chimiques. Ces informations permettent de mieux naviguer dans l'espace combinatoire complexe des molécules et d'optimiser simultanément la synthétisabilité ou d'autres propriétés souhaitées.

Work Package 2 : Apprentissage par renforcement. Le **Work Package 2**, quant à lui, est axé sur l'adaptation et l'amélioration des méthodes d'apprentissage par renforcement (*RL*) appliquées à la génération de molécules. Actuellement, des techniques telles que le *Q-Learning* et d'autres variantes de *RL* sont explorées pour optimiser les actions de mutation dans les graphes moléculaires. Les problématiques abordées incluent la gestion des longues séquences d'actions et l'apprentissage par transfert pour traiter des problèmes plus complexes tout en minimisant les ressources computationnelles.

Work Package 3 : synthétisabilité. Le **Work Package 3**, en lien avec les travaux de thèse de Derin Özer, se concentre sur la génération de molécules en se basant sur des méthodes de synthétisabilité *bottom-up*. Cette approche innovante, par opposition aux méthodes traditionnelles *top-down* de rétrosynthèse, vise à générer des molécules aux propriétés optimales tout en garantissant leur faisabilité en laboratoire. Chaque séquence d'actions est conçue pour respecter les contraintes de synthétisabilité.

Nos recherches actuelles (non publiées) exploitent des modèles de langage de type Transformer, comme T5, pré-entraînés sur des notations *SMILES* et *SMARTS*, afin de sélectionner les réactions chimiques appropriées. Parallèlement, une architecture hiérarchique est employée pour gérer les limites liées au nombre maximal de tokens et permettre le traitement simultané des molécules cibles et des réactions chimiques. Cela garantit non seulement l'optimisation des molécules générées, mais aussi leur synthétisabilité avec des chemins de synthèse clairs et reproductibles.

En complément, notre approche s'appuie sur notre modèle : *ProPreT5* [Oze+25]. *ProPreT5* est une variante optimisée du modèle T5, spécifiquement conçue pour la prédiction de produits de réaction chimique en utilisant un équilibre entre les templates réactionnels très spécifiques et les méthodes sans templates. Ce modèle repose sur l'exploitation du Broad Reaction Set

(BRS), un ensemble de 20 templates réactionnels génériques formulés en notation SMARTS que nous proposons, permettant une exploration chimique plus large que les jeux de données classiques comme USPTO MIT. Contrairement aux approches purement basées sur des brevets, souvent limitées par des réactions spécifiques et une faible capacité de généralisation, notre méthode vise à combiner la flexibilité des transformers avec une contrainte explicite de synthétisabilité. L'intégration de ces outils, que nous avons développés, permet non seulement de conserver une bonne précision des prédictions de réactions chimiques, mais aussi de proposer une solution plus généraliste garantissant des produits réalistes et exploitables en laboratoire.

Perspectives et extensions. Les perspectives de ce travail incluent l'amélioration continue des algorithmes de génération de molécules, rendant le processus plus efficace et applicable à des domaines encore plus complexes comme la chimie verte (*Green Chemistry*) ou la conception de matériaux de haute performance. L'optimisation des algorithmes sera incontournable pour garantir des solutions robustes, réalistes et bien entendu applicables à un niveau industriel.

L'optimisation multi-objectifs, via des méthodes évolutionnaires combinées à l'apprentissage par renforcement, constitue également un autre axe de mes recherches futures. En poursuivant le travail amorcé avec des projets tels qu'EvoMol et G^- GENOCOD, mon objectif est de développer des algorithmes capables d'explorer efficacement de vastes espaces de recherche tout en équilibrant des objectifs complexes, souvent conflictuels, comme la stabilité, la synthétisabilité, et les propriétés fonctionnelles des molécules. Cet axe n'est pas directement intégré dans l'un des Work Packages de G^- GENOCOD ou de la thèse de Dérin Özer, mais il devrait rapidement s'imposer.

5.2.2.2 Organisation du projet G^- GENOCOD

Le projet G^- GENOCOD est prévu pour une durée de 24 mois (du 01/01/2024 au 31/12/2025). Il implique trois *Work Packages (WP)* et cinq chercheurs incluant une post-doctorante durant 18 mois (voir Tableau 5.1).

Quatre mois après le dépôt de la réponse à AAP Étoiles Montantes de ce projet, une doctorante, Dérin Özer, a été recrutée pour traiter plus spécifiquement de la synthétisabilité (Work Package 3 du projet). Sa thèse a commencé en octobre 2024 et encourra jusqu'à fin 2026. Ceci étend donc nos travaux du projet G^- GENOCOD prévus pour deux ans sur une année supplémentaire.

5.2.2.3 Impacts et retombées du projet

Le projet G^- GENOCOD s'inscrit dans la lignée d'approches émergentes des cinq dernières années, exploitant l'Intelligence Artificielle (IA) pour la génération de molécules à des fins de découverte, comme celle de nouveaux médicaments ou de nouveaux matériaux. G^- GENOCOD

Laboratoire	Ressource	Position actuelle	Rôle & responsabilités dans le projet
LERIA	Nicolas Gutowski	Maître de conférences	Coordinateur du projet et Responsable scientifique
LERIA	Benoit Da Mota	Maître de conférences	Co-responsable scientifique
LERIA	Sylvain Lamprier	Professeur des Universités	Conseiller scientifique en science du numérique
MOLTECH-Anjou CNRS UMR 6200	Thomas Cauchy	Maître de conférences	Conseiller scientifique en chimie
LERIA	Gaëlle Milon-Harnois	Post-doctorante	Réalisation opérationnelle et de recherche
LERIA	Jérémy Sanna	Stagiaire M2 Recherche	Réalisation opérationnelle sur les Work Packages 1 et 2

TABLE 5.1 – Tableau récapitulatif des personnes impliquées dans le projet

représente une opportunité d'innovation pour les industries chimiques et des énergies renouvelables, notamment dans des domaines comme la conception de panneaux solaires organiques.

Dans le contexte actuel, marqué par l'essor de modèles génératifs de grande ampleur tels que *GPT*, *Claude*, *Mistral*, *Llama*, *Gemini*, le projet ^G-*GENOCOD* ambitionne de contribuer à la recherche sur les méthodes génératives assistées par IA, mais avec une spécialisation dans la chimie moléculaire. Bien que des géants tels qu'AstraZeneca, IBM et Amazon se soient également engagés dans la course à la génération moléculaire, l'équipe d'origine a su maintenir sa position compétitive depuis le premier projet *EvoMol*. Avec ^G-*GENOCOD*, l'équipe est renforcée et nous visons à consolider notre présence au sein des leaders de ce domaine, avec des perspectives de résultats prometteuses à long terme.

Le projet repose sur une expertise internationalement reconnue en IA et en chimie quantique, grâce aux collaborations entre le *LERIA* et *MOLTECH-Anjou*. Ainsi, ^G-*GENOCOD* se positionne non seulement comme un projet de recherche scientifique, mais aussi comme un vecteur de diffusion pédagogique pour sensibiliser le grand public à l'impact de l'IA et de la chimie moléculaire sur la société.

Sur le plan scientifique, les ambitions du projet sont de taille. ^G-*GENOCOD* aborde un problème fondamental de génération de structures graphiques complexes, appliqué à la chimie, mais qui trouve des résonances dans d'autres domaines de l'IA. La génération de molécules via des choix séquentiels d'actions complexes constitue un enjeu important pour les communautés de l'apprentissage par renforcement et de la chimie moléculaire. En outre, la capture de variables contextuelles riches pour chaque itération représente un obstacle majeur à surmonter, tout comme la synthétisabilité des molécules générées, afin d'assurer que ces solutions puissent être appliquées dans des contextes réels.

Les résultats attendus du projet ouvriront de nouvelles perspectives pour la création de molécules efficaces et spécifiques, et permettront de faire progresser la recherche dans le

domaine de la génération moléculaire et dans la résolution de problèmes complexes de génération séquentielle.

Aujourd'hui, nous prévoyons de publier nos résultats dans des revues internationales en chemo-informatique ou en Intelligence Artificielle. De plus, des conférences de haut niveau dans le domaine de l'intelligence artificielle ainsi que des conférences spécialisées sont actuellement envisagées pour valoriser nos contributions. Des travaux sont en cours d'écriture pour de prochaines soumissions.

Enfin, ^G-GENOCOD servira de base pour l'organisation d'un séminaire national à Angers fin 2025. Nous réfléchissons aujourd'hui à organiser ce séminaire en collaboration avec des groupes de travail reconnus dans le domaine de la génération moléculaire. Cette initiative renforcera non seulement la visibilité du projet mais offrira aussi des opportunités de collaborations et de financements futurs, comme un projet ERC (*European Research Council*).

5.2.3 Perspectives générales et projets de recherche futurs

À court terme, je prévois de soumettre un projet ERC (*European Research Council*) dans la continuité de ^G-GENOCOD, visant à généraliser les méthodes d'apprentissage par renforcement appliquées à la génération de graphes, en particulier dans des environnements multi-objectifs. Ce projet pourra également intégrer des collaborations internationales, notamment dans les domaines de la chimie quantique et de la photonique.

Je souhaite également explorer des projets transdisciplinaires avec des applications dans des domaines aussi variés que la modélisation de systèmes autosuffisants, la simulation de réseaux complexes, et l'optimisation dans des environnements non linéaires avec peu de données disponibles.

Enfin, avec la collaboration en cours avec l'INRIA et la DGA sur l'évaluation des comportements dynamiques non linéaires appliqués aux véhicules terrestres, qui prendra fin en septembre 2025, je prévois de poursuivre ce partenariat avec le même consortium. L'objectif sera de se concentrer sur les capteurs virtuels ou la modélisation du roulis, ce qui me permettra de renforcer mon ancrage pluridisciplinaire, en incluant les compétences requises dans mon département d'affectation, le GEII, au sein de ma composante (IUT).

5.2.4 Conclusion

Les perspectives de recherche que je souhaite poursuivre dans les prochaines années reposent sur des fondations solides issues de mes travaux actuels que je commence à mettre en oeuvre plus récemment dans le cadre du projet ^G-GENOCOD. L'exploration des méthodes d'apprentissage par renforcement, de l'optimisation multi-objectifs et des modèles informés par la physique ouvrent la voie à de nombreuses applications, avec des répercussions potentielles dans divers secteurs scientifiques et industriels. Mon travail continuera à s'inscrire dans une démarche interdisciplinaire et appliquée, avec pour objectif principal d'étendre les frontières des connaissances dans mon domaine de recherche en informatique, en tant que Maître de conférences rattaché à la section 27 du CNU (Conseil national des universités).

BIBLIOGRAPHIE

- [AD01] Claude ABROMONT et Eugène DE MONTALEMBERT, *Guide de la théorie de la musique*, Fayard, 2001.
- [AK17] Laith Mohammad ABUALIGAH et Ahamad Tajudin KHADER, « Unsupervised text feature selection technique based on hybrid particle swarm optimization algorithm with genetic operators for the text clustering », in : *The Journal of Supercomputing* 73 (2017), p. 4773-4795.
- [AT11] Gediminas ADOMAVICIUS et Alexander TUZHILIN, « Context-aware recommender systems », in : *Recommender systems handbook*, Springer, 2011, p. 217-253.
- [AG12] Shipra AGRAWAL et Navin GOYAL, « Analysis of Thompson Sampling for the Multi-armed Bandit Problem », in : *COLT 2012 - The 25th Annual Conference on Learning Theory, June 25-27, 2012, Edinburgh, Scotland*, sous la dir. de Shie MANNOR, Nathan SREBRO et Robert C. WILLIAMSON, t. 23, JMLR Proceedings, JMLR.org, 2012, p. 39.1-39.26.
- [AG13a] Shipra AGRAWAL et Navin GOYAL, « Further Optimal Regret Bounds for Thompson Sampling », in : *Proceedings of the Sixteenth International Conference on Artificial Intelligence and Statistics, AISTATS 2013, Scottsdale, AZ, USA, April 29 - May 1, 2013*, t. 31, JMLR Workshop and Conference Proceedings, JMLR.org, 2013, p. 99-107.
- [AG13b] Shipra AGRAWAL et Navin GOYAL, « Thompson Sampling for Contextual Bandits with Linear Payoffs », in : *Proceedings of the 30th International Conference on Machine Learning, ICML 2013, Atlanta, GA, USA, 16-21 June 2013*, t. 28, JMLR Workshop and Conference Proceedings, JMLR.org, 2013, p. 127-135.
- [Ake+15] Joël AKERET et al., « Approximate Bayesian computation for forward modeling in cosmology », in : *Journal of Cosmology and Astroparticle Physics 2015.08* (août 2015), p. 043.
- [Ale+08] A. ALESSANDRI et al., « Moving-horizon state estimation for nonlinear systems using neural networks », in : *47th IEEE Conference on Decision and Control, 2008*, p. 2557-2562, DOI : 10.1109/CDC.2008.4739462.
- [AF17] Robin ALLESIARDO et Raphaël FÉRAUD, « Selection of learning experts », in : *International Joint Conference on Neural Networks (IJCNN)*, IEEE, 2017, p. 1005-1010, DOI : 10.1109/IJCNN.2017.7965962.
- [Alo+17] Osama Ahmad ALOMARI et al., « Gene selection for cancer classification by combining minimum redundancy maximum relevancy and bat-inspired algorithm », in : *International Journal of Data Mining and Bioinformatics* 19.1 (2017), p. 32-51.

-
- [AK98] Edoardo AMALDI et Viggo KANN, « On the approximability of minimizing nonzero variables or unsatisfied relations in linear systems », in : *Theoretical Computer Science* 209.1-2 (1998), p. 237-260.
- [AVW87] V. ANANTHARAM, P. VARAIYA et J. WALRAND, « Asymptotically efficient allocation rules for the multiarmed bandit problem with multiple plays-Part I : I.I.D. rewards », in : *IEEE Transactions on Automatic Control* 32.11 (1987), p. 968-976, ISSN : 0018-9286, DOI : 10.1109/tac.1987.1104491.
- [AH24] Naveed ANJUM et Ji-Huan HE, « A dual Lagrange multiplier approach for the dynamics of the mechanical systems », in : *Journal of Applied and Computational Mechanics* (2024).
- [AEK00] Asim ANSARI, Skander ESSEGAIER et Rajeev KOHLI, *Internet recommendation systems*, 2000.
- [AP98] Uri M. ASCHER et Linda R. PETZOLD, *Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations*, 1st, USA : Society for Industrial et Applied Mathematics, 1998, ISBN : 0898714125.
- [AO03] Alessandro ASTOLFI et Romeo ORTEGA, « Immersion and invariance : A new tool for stabilization and adaptive control of nonlinear systems », in : *IEEE Transactions on Automatic control* 48.4 (2003), p. 590-606.
- [ABL11] Jean-Yves AUDIBERT, Sébastien BUBECK et Gábor LUGOSI, « Minimax Policies for Combinatorial Prediction Games », in : *Proceedings of the 24th Annual Conference on Learning Theory*, sous la dir. de Sham M. KAKADE et Ulrike von LUXBURG, t. 19, Proceedings of Machine Learning Research, Budapest, Hungary : PMLR, 2011, p. 107-132, URL : <https://proceedings.mlr.press/v19/audibert11a.html>.
- [Aue02] Peter AUER, « Using Confidence Bounds for Exploitation-Exploration Trade-offs », in : *J. Mach. Learn. Res.* 3 (2002), p. 397-422.
- [Aue+95] Peter AUER et al., « Gambling in a rigged casino : The adversarial multi-armed bandit problem », in : *focs*, IEEE, 1995, p. 322.
- [Aue+00] Peter AUER et al., « Gambling in a rigged casino : The adversarial multi-armed bandit problem », in : *Electron. Colloquium Comput. Complex.* TR00-068 (2000), ECCO : TR00-068.
- [Aue+02] Peter AUER et al., « The Nonstochastic Multiarmed Bandit Problem », in : *SIAM J. Comput.* 32.1 (2002), p. 48-77, DOI : 10.1137/S0097539701398375.
- [BZ11] Johannes BADER et Eckart ZITZLER, « HypE : An algorithm for fast hypervolume-based many-objective optimization », in : *Evolutionary computation* 19.1 (2011), p. 45-76.
- [BW14] David BARBER et Yali WANG, « Gaussian Processes for Bayesian Estimation in Ordinary Differential Equations », in : *International Conference on Machine Learning*, 2014.

-
- [BB07] Matthieu BASSEUR et Edmund K BURKE, « Indicator-based multi-objective local search », in : *2007 IEEE Congress on Evolutionary Computation*, IEEE, 2007, p. 3100-3107.
- [Bay+17] Atılım Günes BAYDIN et al., « Automatic differentiation in machine learning : a survey », in : *J. Mach. Learn. Res.* 18.1 (jan. 2017), p. 5595-5637, ISSN : 1532-4435.
- [BGC17] Yoshua BENGIO, Ian GOODFELLOW et Aaron COURVILLE, *Deep learning*, t. 1, MIT press Cambridge, MA, USA, 2017.
- [Ben+23] Laurentino BENITO-EPIGMENIO et al., « Feature Selection : Traditional and Wrapping Techniques with Tabu Search », in : *Innovations in Machine and Deep Learning : Case Studies and Applications*, sous la dir. de Gilberto RIVERA et al., Cham : Springer Nature Switzerland, 2023, p. 21-38, ISBN : 978-3-031-40688-1, DOI : 10.1007/978-3-031-40688-1_2, URL : https://doi.org/10.1007/978-3-031-40688-1_2.
- [BAA23] Hachemi BENNACEUR, Meznah ALMUTAIRY et Norah ALHUSSAIN, « Genetic Algorithm Combined with the K-Means Algorithm : A Hybrid Technique for Unsupervised Feature Selection. », in : *Intelligent Automation & Soft Computing* 37.3 (2023).
- [Ben17] Idir BENOURET, « Un système de recommandation contextuel et composite pour la visite personnalisée de sites culturels », Theses, Université de Technologie de Compiègne, 2017, URL : <https://tel.archives-ouvertes.fr/tel-01767997>.
- [BNE07] Nicola BEUME, Boris NAUJOKS et Michael EMMERICH, « SMS-EMOA : Multiobjective selection based on dominated hypervolume », in : *European Journal of Operational Research* 181.3 (2007), p. 1653-1669.
- [Bic+12] G Richard BICKERTON et al., « Quantifying the chemical beauty of drugs », in : *Nature chemistry* 4.2 (2012), p. 90-98.
- [BDB86] L. T. BIEGLER, J. J. DAMIANO et G. E. BLAU, « Nonlinear parameter estimation : A case study comparison », in : *AIChE Journal* 32.1 (1986), p. 29-45, DOI : <https://doi.org/10.1002/aic.690320105>, eprint : <https://aiche.onlinelibrary.wiley.com/doi/pdf/10.1002/aic.690320105>, URL : <https://aiche.onlinelibrary.wiley.com/doi/abs/10.1002/aic.690320105>.
- [BD20] Kaj BOSTROM et Greg DURRETT, « Byte Pair Encoding is Suboptimal for Language Model Pretraining », in : *Findings of the Association for Computational Linguistics : EMNLP 2020*, Online : Association for Computational Linguistics, nov. 2020, p. 4617-4624, DOI : 10.18653/v1/2020.findings-emnlp.414, URL : <https://aclanthology.org/2020.findings-emnlp.414>.
- [BF16] Djallel BOUNEFFOUF et Raphael FERAUD, « Multi-armed bandit problem with known trend », in : *Neurocomputing* 205 (2016), p. 16-21.

-
- [Bou20] Nadjat BOURDACHE, « Élicitation incrémentale des préférences pour l'optimisation multi-objectifs : modèles non-linéaires, domaines combinatoires et approches tolérantes aux erreurs », Theses, Sorbonne Université, déc. 2020.
- [Bou+18] Konstantinos BOUSMALIS et al., « Using Simulation and Domain Adaptation to Improve Efficiency of Deep Robotic Grasping », in : *2018 IEEE International Conference on Robotics and Automation (ICRA)*, 2018, p. 4243-4250.
- [BK87] Frederick BROOKS et H KUGLER, *No silver bullet*, April, 1987.
- [BO14] Jenný BRYNJARSDÓTTIR et Anthony OHAGAN, « Learning about physical parameters : the importance of model discrepancy », in : *Inverse Problems* 30.11 (oct. 2014), p. 114007.
- [Bur02] Robin D. BURKE, « Hybrid Recommender Systems : Survey and Experiments », in : *User Model. User Adapt. Interact.* 12.4 (2002), p. 331-370, DOI : 10.1023/A:1021240730564.
- [CGL08] Ben CALDERHEAD, Mark GIROLAMI et Neil LAWRENCE, « Accelerating Bayesian Inference over Nonlinear Differential Equations with Gaussian Processes », in : *Advances in Neural Information Processing Systems*, sous la dir. de D. KOLLER et al., t. 21, Curran Associates, Inc., 2008, URL : https://proceedings.neurips.cc/paper_files/paper/2008/file/07563a3fe3bbe7e3ba84431ad9d055af-Paper.pdf.
- [CG92] George CASELLA et Edward I. GEORGE, « Explaining the Gibbs Sampler », in : *The American Statistician* 46.3 (1992), p. 167-174, ISSN : 00031305, DOI : 10.2307/2685208, URL : <http://dx.doi.org/10.2307/2685208>.
- [CLD23] Thomas CAUCHY, Jules LEGUY et Benoit DA MOTA, « Definition and exploration of realistic chemical spaces using the connectivity and cyclic features of ChEMBL and ZINC », in : *Digital Discovery* 2.3 (2023), p. 736-747.
- [Cha+23] Ankush CHAKRABARTY et al., « Moving Horizon Estimation for Digital Twins using Deep Autoencoders », in : *IFAC-PapersOnLine* 56.2 (2023), 22nd IFAC World Congress, p. 5500-5505, ISSN : 2405-8963, DOI : <https://doi.org/10.1016/j.ifacol.2023.10.207>.
- [CH08] Vira CHANKONG et Yacov Y HAIMES, *Multiobjective decision making : theory and methodology*, Courier Dover Publications, 2008.
- [CLL23] Hao CHEN, Shanhe LOU et Chen LV, « Hybrid physics-data-driven online modeling : Framework, methodology and application to electric vehicles », in : *Mechanical Systems and Signal Processing* 185 (2023), p. 109791.
- [CWY13] Wei CHEN, Yajun WANG et Yang YUAN, « Combinatorial Multi-Armed Bandit : General Framework and Applications », in : *Proceedings of the 30th International Conference on Machine Learning, ICML 2013, Atlanta, GA, USA, 16-21 June 2013*, t. 28, JMLR Workshop and Conference Proceedings, JMLR.org, 2013, p. 151-159.

-
- [CX21] Zhen CHEN et Dongbin XIU, « On generalized residual network for deep learning of unknown dynamical systems », in : *Journal of Computational Physics* 438 (2021), p. 110362, ISSN : 0021-9991.
- [CL01] Eddie WL CHENG et Heng LI, « Information priority-setting for better resource allocation using analytic hierarchy process (AHP) », in : *Information Management & Computer Security* 9.2 (2001), p. 61-70.
- [CG95] Siddhartha CHIB et Edward GREENBERG, « Understanding the Metropolis-Hastings Algorithm », in : *The American Statistician* 49.4 (1995), p. 327-335, ISSN : 00031305, URL : <http://www.jstor.org/stable/2684568> (visité le 30/05/2023).
- [CTJ21] Davide CHICCO, Niklas TÖTSCH et Giuseppe JURMAN, « The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation », in : *BioData mining* 14.1 (2021), p. 1-22.
- [Coe07] CAC COELLO, *Evolutionary Algorithms for solving Multi-Objective Problems*, 2007.
- [Coe99] Carlos A COELLO COELLO, « A comprehensive survey of evolutionary-based multiobjective optimization techniques », in : *Knowledge and Information systems* 1.3 (1999), p. 269-308.
- [Com+15] Richard COMBES et al., « Combinatorial Bandits Revisited », in : *Advances in Neural Information Processing Systems 28 : Annual Conference on Neural Information Processing Systems 2015, December 7-12, 2015, Montreal, Quebec, Canada*, sous la dir. de Corinna CORTES et al., 2015, p. 2116-2124.
- [Cos+24] Raphaël COSSON et al., « On bi-objective combinatorial optimization with heterogeneous objectives », in : *European Journal of Operational Research* 319.1 (2024), p. 89-101, ISSN : 0377-2217, DOI : <https://doi.org/10.1016/j.ejor.2024.06.029>, URL : <https://www.sciencedirect.com/science/article/pii/S0377221724004776>.
- [Cra+20] Miles CRANMER et al., *Lagrangian Neural Networks*, 2020, arXiv : 2003.04630 [cs.LG].
- [Deb11] Kalyanmoy DEB, « Multi-objective optimisation using evolutionary algorithms : an introduction », in : *Multi-objective evolutionary optimisation for product design and manufacturing*, Springer, 2011, p. 3-34.
- [Deb+02] Kalyanmoy DEB et al., « A fast and elitist multiobjective genetic algorithm : NSGA-II », in : *IEEE transactions on evolutionary computation* 6.2 (2002), p. 182-197.
- [DK04] Mukund DESHPANDE et George KARYPIS, « Item-based top-n recommendation algorithms », in : *ACM Transactions on Information Systems (TOIS)* 22.1 (2004), p. 143-177.

-
- [Don+19] Chris DONAHUE et al., « LakhNES : Improving Multi-instrumental Music Generation with Cross-domain Pre-training », in : *Proceedings of the 20th International Society for Music Information Retrieval Conference, ISMIR 2019, Delft, The Netherlands, November 4-8, 2019*, 2019, p. 685-692, URL : <http://archives.ismir.net/ismir2019/paper/000083.pdf>.
- [Don+18] Hao-Wen DONG et al., « MuseGAN : Multi-track Sequential Generative Adversarial Networks for Symbolic Music Generation and Accompaniment », in : *Proceedings of the AAAI Conference on Artificial Intelligence 32.1* (avr. 2018), URL : <https://ojs.aaai.org/index.php/AAAI/article/view/11312>.
- [Don+23] Hao-Wen DONG et al., « Multitrack Music Transformer », in : *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2023, p. 1-5, DOI : 10.1109/ICASSP49357.2023.10094628, URL : <https://ieeexplore.ieee.org/document/10094628>.
- [DN13] Madalina M. DRUGAN et Ann NOWÉ, « Designing multi-objective multi-armed bandits algorithms : A study », in : *The 2013 International Joint Conference on Neural Networks, IJCNN 2013, Dallas, TX, USA, August 4-9, 2013*, IEEE, 2013, p. 1-8, DOI : 10.1109/IJCNN.2013.6707036.
- [Eft+02] A EFTAXIAS et al., « Nonlinear kinetic parameter estimation using simulated annealing », in : *Computers Chemical Engineering* 26.12 (2002), p. 1725-1733, ISSN : 0098-1354, DOI : [https://doi.org/10.1016/S0098-1354\(02\)00156-4](https://doi.org/10.1016/S0098-1354(02)00156-4), URL : <https://www.sciencedirect.com/science/article/pii/S0098135402001564>.
- [EF09] Mohsen EJDAY et Lionel FOURMENT, « Optimisation multi-objectifs à base de méta-modèle pour les problèmes de mise en forme », in : *9ème Colloque National en Calcul des Structures*, MS3 : Optimisation multi-disciplinaire, LMT-Cachan, 2009, 6-pages.
- [ERR16] Mehdi ELAHI, Francesco RICCI et Neil RUBENS, « A survey of active learning in collaborative filtering recommender systems », in : *Computer Science Review* 20 (juin 2016), DOI : 10.1016/j.cosrev.2016.05.002.
- [Ela+13] Mehdi ELAHI et al., « Personality-Based Active Learning for Collaborative Filtering Recommender Systems », in : *Advances in Artificial Intelligence*, sous la dir. de Matteo BALDONI et al., Springer International Publishing, 2013, p. 360-371, ISBN : 978-3-319-03524-6.
- [EP20] Jeff ENS et Philippe PASQUIER, *MMM : Exploring Conditional Multi-Track Music Generation with the Transformer*, 2020, arXiv : 2008.06048 [cs.SD], URL : <https://arxiv.org/abs/2008.06048>.
- [ES09] Peter ERTL et Ansgar SCHUFFENHAUER, « Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions », in : *Journal of cheminformatics* 1 (2009), p. 1-11.

-
- [Esf+23] Hossein Nejatbakhsh ESFAHANI et al., « Learning-based state estimation and control using MHE and MPC schemes with imperfect models », in : *European Journal of Control* 73 (2023), p. 100880.
- [FT23] Gianni FILOGRASSO et Giacomo di TOLLO, « Adaptive evolutionary algorithms for portfolio selection problems », in : *Computational Management Science* 20.1 (2023), p. 7.
- [Fra24] Nathan FRADET, « Apprentissage automatique pour la modélisation de musique symbolique », 2024SORUS037, thèse de doct., 2024, URL : <http://www.theses.fr/2024SORUS037/document>.
- [Fra+21] Nathan FRADET et al., *MidiTok : A Python package for MIDI file tokenization*, nov. 2021, URL : <https://github.com/Natooz/MidiTok>.
- [Fra+23a] Nathan FRADET et al., « Byte Pair Encoding for Symbolic Music », in : *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, sous la dir. d'Houda BOUAMOR, Juan PINO et Kalika BALI, Singapore : Association for Computational Linguistics, déc. 2023, p. 2001-2020, DOI : 10.18653/v1/2023.emnlp-main.123, URL : <https://aclanthology.org/2023.emnlp-main.123>.
- [Fra+23b] Nathan FRADET et al., « Impact of time and note duration tokenizations on deep learning symbolic music modeling », in : *Proceedings of the 24th International Society for Music Information Retrieval Conference*, Milano, Italy, nov. 2023, URL : <https://hal.science/hal-04147659>.
- [Gab+11] Victor GABILLON et al., « Multi-Bandit Best Arm Identification », in : *Advances in Neural Information Processing Systems 24 : 25th Annual Conference on Neural Information Processing Systems 2011. Proceedings of a meeting held 12-14 December 2011, Granada, Spain*, sous la dir. de John SHAWE-TAYLOR et al., 2011, p. 2222-2230.
- [GSP16] Bhargav GADHVI, Vimal SAVSANI et Vivek PATEL, « Multi-objective optimization of vehicle passive suspension system using NSGA-II, SPEA2 and PESA-II », in : *Procedia Technology* 23 (2016), p. 361-368.
- [Gag94] Philip GAGE, « A New Algorithm for Data Compression », in : *C Users J.* 12.2 (fév. 1994), p. 23-38, ISSN : 0898-9788, URL : <https://dl.acm.org/doi/10.5555/177910.177914>.
- [GS06] Matteo GAGLIOLO et Jürgen SCHMIDHUBER, « Learning dynamic algorithm portfolios », in : *Annals of Mathematics and Artificial Intelligence* 47.3-4 (2006), p. 295-328, DOI : 10.1007/s10472-006-9036-z.
- [GS11] Matteo GAGLIOLO et Jürgen SCHMIDHUBER, « Algorithm portfolio selection as a bandit problem with unbounded losses », in : *Ann. Math. Artif. Intell.* 61.2 (2011), p. 49-86, DOI : 10.1007/S10472-011-9228-Z.

-
- [GYC21] Tianyu GAO, Xingcheng YAO et Danqi CHEN, « SimCSE : Simple Contrastive Learning of Sentence Embeddings », in : *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, Online et Punta Cana, Dominican Republic : Association for Computational Linguistics, nov. 2021, p. 6894-6910, DOI : 10.18653/v1/2021.emnlp-main.552, URL : <https://aclanthology.org/2021.emnlp-main.552>.
- [GR98] Zoubin GHAHRAMANI et Sam ROWEIS, « Learning Nonlinear Dynamical Systems Using an EM Algorithm », in : *Advances in Neural Information Processing Systems*, sous la dir. de M. KEARNS, S. SOLLA et D. COHN, t. 11, MIT Press, 1998.
- [Gol89] David E GOLBERG, « Genetic algorithms in search, optimization, and machine learning », in : *Addion wesley 1989.102* (1989), p. 36.
- [Goo16] Ian GOODFELLOW, *Deep learning*, 2016.
- [GDY19] Samuel GREYDANUS, Misko DZAMBA et Jason YOSINSKI, « Hamiltonian Neural Networks », in : *Advances in Neural Information Processing Systems*, sous la dir. de H. WALLACH et al., t. 32, Curran Associates, Inc., 2019.
- [Gun18] Nyoman GUNANTARA, « A review of multi-objective optimization : Methods and its applications », in : *Cogent Engineering 5.1* (2018), sous la dir. de Qingsong AI, p. 1502242, DOI : 10.1080/23311916.2018.1502242.
- [Guo+21] Jinglei GUO et al., « A Niche Based Multi-objective Particle Swarm Optimizer », in : *2021 IEEE Congress on Evolutionary Computation (CEC)*, IEEE, 2021, p. 1319-1326.
- [Gup+21] Samarth GUPTA et al., « Multi-Armed Bandits With Correlated Arms », in : *IEEE Transactions on Information Theory 67.10* (2021), p. 6711-6732, ISSN : 1557-9654, DOI : 10.1109/tit.2021.3081508, URL : <http://dx.doi.org/10.1109/TIT.2021.3081508>.
- [Gut19] Nicolas GUTOWSKI, « Recommandation contextuelle de services : Application à la recommandation d'événements culturels dans la ville intelligente. (Context-aware recommendation systems for cultural events recommendation in Smart Cities) », thèse de doct., University of Angers, France, 2019.
- [Gut+17] Nicolas GUTOWSKI et al., « A Framework for Context-Aware Service Recommendation for Mobile Users : A Focus on Mobility in Smart Cities », in : *From Data To Decision* (2017).
- [Gut+18] Nicolas GUTOWSKI et al., *Context Enhancement for Linear Contextual Multi-Armed Bandits*, 2018, DOI : 10.1109/ictai.2018.00161.
- [Gut+19a] Nicolas GUTOWSKI et al., « Global versus individual accuracy in contextual multi-armed bandit », in : *Proceedings of the 34th ACM/SIGAPP symposium on applied computing*, 2019, p. 1647-1654.

-
- [Gut+19b] Nicolas GUTOWSKI et al., « Gorthaur : A Portfolio Approach for Dynamic Selection of Multi-Armed Bandit Algorithms for Recommendation », in : *2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI)*, IEEE, 2019, p. 1164-1171, DOI : 10.1109/ICTAI.2019.00161.
- [Gut+19c] Nicolas GUTOWSKI et al., « Improving Bandit-Based Recommendations with Spatial Context Reasoning : An Online Evaluation », in : *2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI)*, IEEE, 2019, p. 1366-1373.
- [Gut+19d] Nicolas GUTOWSKI et al., « Using individual accuracy to create context for non-contextual multi-armed bandit problems », in : *2019 IEEE-RIVF International Conference on Computing and Communication Technologies (RIVF)*, IEEE, 2019, p. 1-6.
- [Gut+21] Nicolas GUTOWSKI et al., « Gorthaur-EXP3 : Bandit-based selection from a portfolio of recommendation algorithms balancing the accuracy-diversity dilemma », in : *Inf. Sci.* 546 (2021), p. 378-396, DOI : 10.1016/J.INS.2020.08.106.
- [Gut+22] Nicolas GUTOWSKI et al., « A novel multi-objective medical feature selection compass method for binary classification », in : *Artif. Intell. Medicine* 127 (2022), p. 102277, DOI : 10.1016/J.ARTMED.2022.102277.
- [Haa+16] Tuomas HAARNOJA et al., « Backprop KF : Learning Discriminative Deterministic State Estimators », in : *Advances in Neural Information Processing Systems*, sous la dir. de D. LEE et al., t. 29, Curran Associates, Inc., 2016.
- [HC21] Gaëtan HADJERES et Léopold CRESTEL, *The Piano Inpainting Application*, 2021, arXiv : 2107.05944 [cs.SD], URL : <https://arxiv.org/abs/2107.05944>.
- [HPN17] Gaëtan HADJERES, François PACHET et Frank NIELSEN, « DeepBach : a Steerable Model for Bach Chorales Generation », in : *Proceedings of the 34th International Conference on Machine Learning*, sous la dir. de Doina PRECUP et Yee Whye TEH, t. 70, Proceedings of Machine Learning Research, PMLR, août 2017, p. 1362-1371, URL : <https://proceedings.mlr.press/v70/hadjeres17a.html>.
- [Ham+07] Tarek M HAMDANI et al., « Multi-objective feature selection with NSGA II », in : *Adaptive and Natural Computing Algorithms : 8th International Conference, ICANNGA 2007, Warsaw, Poland, April 11-14, 2007, Proceedings, Part I* 8, Springer, 2007, p. 240-247.
- [Han22] Emrah HANCER, « A multi-objective Artificial Bee Colony algorithm for cost-sensitive subset selection », in : *Neural Computing and Applications* 34.20 (2022), p. 17523-17537.
- [Han16] Nikolaus HANSEN, « The CMA evolution strategy : A tutorial », in : *arXiv preprint arXiv :1604.00772* (2016).

-
- [He+16] Kaiming HE et al., « Deep Residual Learning for Image Recognition », in : *Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR '16, Las Vegas, NV, USA* : IEEE, juin 2016, p. 770-778, DOI : 10.1109/CVPR.2016.90.
- [Hir+19] Gaurush HIRANANDANI et al., « Cascading Linear Submodular Bandits : Accounting for Position Bias and Diversity in Online Learning to Rank », in : *Proceedings of the Thirty-Fifth Conference on Uncertainty in Artificial Intelligence, UAI 2019, Tel Aviv, Israel, July 22-25, 2019*, sous la dir. d'Amir GLOBERSON et Ricardo SILVA, t. 115, Proceedings of Machine Learning Research, AUAI Press, 2019, p. 722-732.
- [HP02] Yu-Chi HO et David L PEPYNE, « Simple explanation of the no-free-lunch theorem and its implications », in : *Journal of optimization theory and applications* 115.3 (2002), p. 549-570, DOI : 10.1109/CDC.2001.980896.
- [HP01] Yu-Chi HO et David L. PEPYNE, « Simple explanation of the no free lunch theorem of optimization », in : *40th IEEE Conference on Decision and Control, CDC 2001, Orlando, FL, USA, 4-7 Dec., 2001*, IEEE, 2001, p. 4409-4414, DOI : 10.1109/.2001.980896.
- [Hoc97] S HOCHREITER, « Long Short-term Memory », in : *Neural Computation MIT-Press* (1997).
- [Hol92] John H. HOLLAND, *Adaptation in Natural and Artificial Systems : An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*, The MIT Press, avr. 1992, ISBN : 9780262275552.
- [Hsi+21] Wen-Yi HSIAO et al., « Compound Word Transformer : Learning to Compose Full-Song Music over Dynamic Directed Hypergraphs », in : *Proceedings of the AAAI Conference on Artificial Intelligence* 35.1 (mai 2021), p. 178-186, DOI : 10.1609/aaai.v35i1.16091, URL : <https://ojs.aaai.org/index.php/AAAI/article/view/16091>.
- [Hu+21] Weifei HU et al., « Digital Twin : A State-of-the-Art Review of Its Enabling Technologies, Applications and Challenges », in : *Journal of Intelligent Manufacturing and Special Equipment* 2.1 (2021), p. 1-34, ISSN : 2633-6596, 2633-660X.
- [Hua+17] Cheng-Zhi Anna HUANG et al., « Counterpoint by Convolution », in : *Proceedings of 18st International Conference on Music Information Retrieval, ISMIR, 2017*, URL : <https://arxiv.org/abs/1903.07227>.
- [Hua+18] Cheng-Zhi Anna HUANG et al., « Music transformer », in : *arXiv preprint arXiv :1809.04281* (2018).
- [HL16] Kuan-Hao HUANG et Hsuan-Tien LIN, « Linear Upper Confidence Bound Algorithm for Contextual Bandit Problem with Piled Rewards », in : avr. 2016, p. 143-155, ISBN : 978-3-319-31749-6, DOI : 10.1007/978-3-319-31750-2_12.

-
- [HY20] Yu-Siang HUANG et Yi-Hsuan YANG, « Pop Music Transformer : Beat-Based Modeling and Generation of Expressive Pop Piano Compositions », in : *Proceedings of the 28th ACM International Conference on Multimedia*, MM '20, Seattle, WA, USA : Association for Computing Machinery, 2020, p. 1180-1188, ISBN : 9781450379885, DOI : 10.1145/3394171.3413671, URL : <https://doi.org/10.1145/3394171.3413671>.
- [Hun+21] Hsiao-Tzu HUNG et al., « EMOPIA : A Multi-Modal Pop Piano Dataset For Emotion Recognition and Emotion-based Music Generation », in : *Proceedings of the 22nd International Society for Music Information Retrieval Conference, ISMIR 2021, Online, November 7-12, 2021*, sous la dir. de Jin Ha LEE et al., 2021, p. 318-325, URL : <https://archives.ismir.net/ismir2021/paper/000039.pdf>.
- [IBK06] Edward L IONIDES, Carles BRETÓ et Aaron A KING, « Inference for nonlinear dynamical systems », in : *Proceedings of the National Academy of Sciences* 103.49 (2006), p. 18438-18443.
- [Ito+19] Shinji ITO et al., « Improved Regret Bounds for Bandit Combinatorial Optimization », in : *Advances in Neural Information Processing Systems 32 : Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8-14, 2019, Vancouver, BC, Canada*, sous la dir. d'Hanna M. WALLACH et al., 2019, p. 12027-12036.
- [Jai22] Shashank Mohan JAIN, « Hugging face », in : *Introduction to transformers for NLP : With the hugging face library and models to solve problems*, Springer, 2022, p. 51-67.
- [JF13] Dietmar JANNACH et Gerhard FRIEDRICH, « Tutorial : recommender systems », in : *International Joint Conference on Artificial Intelligence Beijing*, 2013.
- [Jeb+07] Mohamed JEBALIA et al., « Identification of the isotherm function in chromatography using CMA-ES », in : *2007 IEEE Congress on Evolutionary Computation*, IEEE, 2007, p. 4289-4296.
- [Jia+23] Ruwang JIAO et al., « A survey on evolutionary multiobjective feature selection in classification : approaches, applications, and challenges », in : *IEEE Transactions on Evolutionary Computation* (2023).
- [Jim+17] Magdiel JIMÉNEZ-GUARNEROS et al., « Long-Term Prediction of a Sine Function Using a LSTM Neural Network », in : *Nature-Inspired Design of Hybrid Intelligent Systems*, sous la dir. de Patricia MELIN, Oscar CASTILLO et Janusz KACPRZYK, Cham : Springer International Publishing, 2017, p. 159-173, ISBN : 978-3-319-47054-2.
- [KS17] A KALE et S SONAVANE, « Hybrid feature subset selection approach for fuzzy-extreme learning machine », in : *Data-Enabled Discovery and Applications* 1 (2017), p. 1-14.

-
- [Kal60] Rudolf E. KALMAN, « A New Approach to Linear Filtering and Prediction Problems », in : *Transactions of the ASME—Journal of Basic Engineering* 82.Series D, *Journal of Basic Engineering* (1960), p. 35-45.
- [Kal97] Rudolf E. KALMAN, « Signal Processing, Sensor Fusion, and Target Recognition VI », in : *Proceedings Volume 3068, AeroSense '97*, International Society for Optics et Photonics, Orlando, FL, United States, 1997, DOI : 10.1117/12.280797.
- [KB61] Rudolf E. KALMAN et Robert S. BUCH, « New Results in Linear Filtering and Prediction Theory », in : *Transactions of the ASME Journal of Basic Engineering* 83 (1961), p. 95-108, DOI : 10.1115/1.3658902.
- [Kar+15] Gülşah KARAKAYA et al., « Identifying (quasi) equally informative subsets in feature selection problems for classification : a max-relevance min-redundancy approach », in : *IEEE transactions on cybernetics* 46.6 (2015), p. 1424-1437.
- [Kar82] Ilhami I KARAYALCIN, *The analytic hierarchy process : Planning, priority setting, resource allocation : Thomas L. SAATY McGraw-Hill, New York, 1980, xiii+ 287 pages,£ 15.65*, 1982.
- [KL21] Benjamin KARG et Sergio LUCIA, « Approximate moving horizon estimation and robust nonlinear model predictive control via deep learning », in : *Computers Chemical Engineering* 148 (2021), p. 107266, ISSN : 0098-1354, DOI : <https://doi.org/10.1016/j.compchemeng.2021.107266>.
- [Kar+22] Maryam KARIMI-MAMAGHAN et al., « Machine learning at the service of meta-heuristics for solving combinatorial optimization problems : A state-of-the-art », in : *European Journal of Operational Research* 296.2 (2022), p. 393-422.
- [Kar+21] George Em KARNIADAKIS et al., « Physics-informed machine learning », in : *Nature Reviews Physics* 3.6 (2021), p. 422-440.
- [KN13] Shima KASHEF et Hossein NEZAMABADI-POUR, « A new feature selection algorithm based on binary ant colony optimization », in : *The 5th Conference on Information and Knowledge Technology*, 2013, p. 50-54, DOI : 10.1109/IKT.2013.6620037.
- [Kee93] Ralph L KEENEY, *Decisions with multiple objectives : Preferences and value tradeoffs*, Cambridge university press, 1993.
- [Ker+19] Pascal KERSCHKE et al., « Automated Algorithm Selection : Survey and Perspectives », in : *Evolutionary Computation* 27.1 (2019), p. 3-45, ISSN : 1063-6560, DOI : 10.1162/evco_a_00242.
- [KB15] Diederik P. KINGMA et Jimmy BA, « Adam : A Method for Stochastic Optimization », in : *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, 2015, URL : <http://arxiv.org/abs/1412.6980>.

-
- [KSS] Pushmeet KOHLI, Mahyar SALEK et Greg STODDARD, « A Fast Bandit Algorithm for Recommendation to Users With Heterogenous Tastes », in : *Proceedings of the AAAI Conference on Artificial Intelligence 27.1* (), p. 1135-1141, ISSN : 2374-3468, DOI : 10.1609/aaai.v27i1.8463.
- [Kom+17] Majid KOMEILI et al., « Feature selection for nonstationary data : Application to human recognition using medical biometrics », in : *IEEE transactions on cybernetics* 48.5 (2017), p. 1446-1459.
- [KCW20] Qiuqiang KONG, Keunwoo CHOI et Yuxuan WANG, *Large-Scale MIDI-based Composer Classification*, 2020, arXiv : 2010.14805 [cs.SD], URL : <https://arxiv.org/abs/2010.14805>.
- [Kon+21] Qiuqiang KONG et al., « GiantMIDI-Piano : A large-scale midi dataset for classical piano music », in : *Transactions of the International Society for Music Information Retrieval*, t. 5, 2021, p. 87-98, DOI : 10.5334/tismir.80, URL : <https://transactions.ismir.net/articles/10.5334/tismir.80/#>.
- [Kud18] Taku KUDO, « Subword Regularization : Improving Neural Network Translation Models with Multiple Subword Candidates », in : *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1 : Long Papers)*, Melbourne, Australia : Association for Computational Linguistics, juill. 2018, p. 66-75, DOI : 10.18653/v1/P18-1007, URL : <https://aclanthology.org/P18-1007>.
- [Kue+17] Alex KUEFLER et al., « Imitating driver behavior with generative adversarial networks », in : *2017 IEEE intelligent vehicles symposium (IV)*, IEEE, 2017, p. 204-211.
- [Kve+15] Branislav KVETON et al., « Cascading Bandits : Learning to Rank in the Cascade Model », in : *Proceedings of the 32nd International Conference on Machine Learning*, sous la dir. de Francis BACH et David BLEI, t. 37, Proceedings of Machine Learning Research, Lille, France : PMLR, 2015, p. 767-776.
- [LZ08] John LANGFORD et Tong ZHANG, « The epoch-greedy algorithm for multi-armed bandits with side information », in : *Advances in neural information processing systems*, 2008, p. 817-824.
- [LF18] Romain LAROCHE et Raphaël FÉRAUD, « Reinforcement Learning Algorithm Selection », in : *6th International Conference on Learning Representations, ICLR 2018, Vancouver, BC, Canada, April 30 - May 3, 2018, Conference Track Proceedings*, OpenReview.net, 2018.
- [LS20] Tor LATTIMORE et Csaba SZEPESVÁRI, *Bandit Algorithms*, 2020, DOI : 10.1017/9781108571401.
- [LBH15] Yann LECUN, Yoshua BENGIO et Geoffrey HINTON, « Deep learning », in : *nature* 521.7553 (2015), p. 436-444.

-
- [Led72] John Ronald Reuel Tolkien (trad. Francis LEDOUX), *Le Seigneur des Anneaux*, 1972, p. 67.
- [Leg+20] Jules LEGUY et al., « EvoMol : a flexible and interpretable evolutionary algorithm for unbiased de novo molecular generation », in : *Journal of cheminformatics* 12 (2020), p. 1-19.
- [Leg+21] Jules LEGUY et al., « Scalable estimator of the diversity for de novo molecular generation resulting in a more robust QM dataset (OD9) and a more efficient molecular optimization », in : *Journal of cheminformatics* 13 (2021), p. 1-17.
- [Leg+22] Jules LEGUY et al., « Goal-directed generation of new molecules by AI methods », in : *Computational and Data-Driven Chemistry Using Artificial Intelligence*, Elsevier, 2022, p. 39-67.
- [Let+21] A LETARD et al., « Bandits-Manchots Combinatoires : du retour utilisateur à la recommandation », in : *CNIA 2021 : Conférence Nationale en Intelligence Artificielle*, 2021, p. 52-59.
- [Let23] A. LETARD, « Un système d'apprentissage multicritère guidé par les besoins utilisateur », Theses, Université d'Angers, 2023.
- [Let+20a] Alexandre LETARD et al., « Bandit et Semi-Bandit avec Retour Partiel : Une Stratégie d'Optimisation du Retour Utilisateur », in : *5ème Conférence Nationale sur les Applications Pratiques de l'Intelligence Artificielle (APIA)*, 2020.
- [Let+20b] Alexandre LETARD et al., « Partial Bandit and Semi-Bandit : Making the Most Out of Scarce Users' Feedback », in : *ICTAI*, 2020.
- [Let+22] Alexandre LETARD et al., « COM-MABs : From Users' Feedback to Recommendation », in : *Proceedings of the Thirty-Fifth International Florida Artificial Intelligence Research Society Conference, FLAIRS 2022, Hutchinson Island, Jensen Beach, Florida, USA, May 15-18, 2022*, sous la dir. de Roman BARTÁK, Fazel KESHTKAR et Michael FRANKLIN, 2022, DOI : 10.32473/FLAIRS.V35I.130560.
- [Let+24] Alexandre LETARD et al., « Bandit algorithms : A comprehensive review and their dynamic selection from a portfolio for multicriteria top-k recommendation », in : *Expert Systems with Applications* (2024), p. 123151.
- [Li+24] Cuixia LI et al., « A compass-based hyper-heuristic for multi-objective optimization problems », in : *Swarm and Evolutionary Computation* 87 (2024), p. 101530.
- [Li+20] Haoran LI et al., « A dividing-based many-objective evolutionary algorithm for large-scale feature selection », in : *Soft computing* 24 (2020), p. 6851-6870.
- [Li+10] Lihong LI et al., « A contextual-bandit approach to personalized news article recommendation », in : *Proceedings of the 19th International Conference on World Wide Web, WWW 2010, Raleigh, North Carolina, USA, April 26-30, 2010*, sous la dir. de Michael RAPPA et al., ACM, 2010, p. 661-670, DOI : 10.1145/1772690.1772758.

-
- [Lia+17] Feynman T. LIANG et al., « Automatic Stylistic Composition of Bach Chorales with Deep LSTM », in : *Proceedings of the 18th International Society for Music Information Retrieval Conference, ISMIR 2017, Suzhou, China, October 23-27, 2017*, 2017, p. 449-456, URL : https://ismir2017.smcnus.org/wp-content/uploads/2017/10/156%5C_Paper.pdf.
- [LK16] Linxia LIAO et Felix KÖTTIG, « A hybrid framework combining data-driven and model-based methods for system remaining useful life prediction », in : *Applied Soft Computing* 44 (2016), p. 191-199.
- [LTV24] Arnaud LIEFOOGHE, Ryoji TANABE et Sébastien VEREL, « Contrasting the landscapes of feature selection under different machine learning models », in : *International Conference on Parallel Problem Solving from Nature*, Springer, 2024, p. 360-376.
- [Lie+07] Arnaud LIEFOOGHE et al., « ParadisEO-MOEO : A framework for evolutionary multi-objective optimization », in : *Evolutionary Multi-Criterion Optimization : 4th International Conference, EMO 2007, Matsushima, Japan, March 5-8, 2007. Proceedings 4*, Springer, 2007, p. 386-400.
- [Lie+19] Arnaud LIEFOOGHE et al., « Landscape-aware performance prediction for evolutionary multiobjective optimization », in : *IEEE Transactions on Evolutionary Computation* 24.6 (2019), p. 1063-1077.
- [Lie+14] Juliane LIEPE et al., « A framework for parameter estimation and model selection from experimental data in systems biology using approximate Bayesian computation », in : *Nature Protocols* 9.2 (2014), p. 439-456, ISSN : 1750-2799, DOI : 10.1038/nprot.2014.025, URL : <https://doi.org/10.1038/nprot.2014.025>.
- [LK10] Gabriele LILLACCI et Mustafa KHAMMASH, « Parameter Estimation and Model Selection in Computational Biology », in : *PLOS Computational Biology* 6.3 (mars 2010), p. 1-17, DOI : 10.1371/journal.pcbi.1000696, URL : <https://doi.org/10.1371/journal.pcbi.1000696>.
- [Lio+23] Fabien LIONTI et al., « Approche Bayésienne pour l'Estimation des Paramètres de la Dynamique Latérale d'un Véhicule », in : *Conference on Artificial Intelligence for Defense*, 2023.
- [Lio+24] Fabien LIONTI et al., « Bayesian Approach for Parameter Estimation in Vehicle Lateral Dynamics », in : *Foundations of Intelligent Systems*, sous la dir. d'Annalisa APPICE et al., Cham : Springer Nature Switzerland, 2024, p. 249-259, ISBN : 978-3-031-62700-2.
- [LL12] Adam LIPOWSKI et Dorota LIPOWSKA, « Roulette-wheel selection via stochastic acceptance », in : *Physica A : Statistical Mechanics and its Applications* 391.6 (2012), p. 2193-2196.
- [LM96] M LITTMAN et A MOORE, *Reinforcement learning : A survey, journal of artificial intelligence research* 4, 1996.

-
- [Liu+22] Jiafeng LIU et al., « Symphony Generation with Permutation Invariant Language Model », in : *Proceedings of the 23rd International Society for Music Information Retrieval Conference*, Bengalore, India : ISMIR, déc. 2022, URL : <https://arxiv.org/abs/2205.05448>.
- [Liu+19] Yan-Jun LIU et al., « Adaptive neural network learning controller design for a class of nonlinear systems with time-varying state constraints », in : *IEEE Transactions on Neural Networks and Learning Systems* 31.1 (2019), p. 66-75.
- [Lju99] L. LJUNG, *System Identification : Theory for the User*, Prentice Hall information and system sciences series, Prentice Hall PTR, 1999, ISBN : 9780136566953.
- [Lou16] Jonathan LOUËDEC, « Stratégies de bandit pour les systèmes de recommandation. (Bandit strategies for recommender systems) », thèse de doct., Paul Sabatier University, Toulouse, France, 2016.
- [Lu+15] Jie LU et al., « Recommender system application developments : a survey », in : *Decision Support Systems* 74 (2015), p. 12-32.
- [LKC19] Alex LUEDTKE, Emilie KAUFMANN et Antoine CHAMBAZ, « Asymptotically optimal algorithms for budgeted multiple play bandits », in : *Machine Learning* 108.11 (nov. 2019), p. 1919-1949, ISSN : 1573-0565, DOI : 10.1007/s10994-019-05799-x.
- [Lue64] David G. LUENBERGER, « Observing the State of a Linear System », in : *IEEE Transactions on Military Electronics* 8.2 (1964), p. 74-80, DOI : 10.1109/TME.1964.4323124.
- [MX17] Benteng MA et Yong XIA, « A tribe competition-based genetic algorithm for feature selection in pattern classification », in : *Applied Soft Computing* 58 (2017), p. 328-338.
- [Ma+23] Haiping MA et al., « A comprehensive survey on NSGA-II for multi-objective optimization and applications », in : *Artificial Intelligence Review* 56.12 (2023), p. 15217-15270.
- [Mar+12] Jean-Michel MARIN et al., « Approximate Bayesian computational methods », in : *Statistics and Computing* 22.6 (nov. 2012), DOI : 10.1007/s11222-011-9288-2, URL : <https://hal.science/hal-00567240>.
- [MA10] R Timothy MARLER et Jasbir S ARORA, « The weighted sum method for multi-objective optimization : new insights », in : *Structural and multidisciplinary optimization* 41 (2010), p. 853-862.
- [MS08] Jorge MATURANA et Frédéric SAUBION, « A Compass to Guide Genetic Algorithms », in : *Parallel Problem Solving from Nature - PPSN X, 10th International Conference Dortmund, Germany, September 13-17, 2008, Proceedings*, sous la dir. de Günther RUDOLPH et al., t. 5199, Lecture Notes in Computer Science, Springer, 2008, p. 256-265, DOI : 10.1007/978-3-540-87700-4_26.

-
- [Mav09] George MAVROTAS, « Effective implementation of the ε -constraint method in multi-objective mathematical programming problems », in : *Applied mathematics and computation* 213.2 (2009), p. 455-465.
- [McC+06] Kevin MCCARTHY et al., « Cats : A synchronous approach to collaborative group recommendation », in : *Florida Artificial Intelligence Research Society Conference (FLAIRS)*, 2006, p. 86-91.
- [Mez+19] Tinhinan MEZDAD et al., « Identification of new factors associated to walking impairment in patients with vascular-type claudication », in : *Vascular* 27.3 (2019), p. 260-269.
- [MM95] H. MICHALSKA et D.Q. MAYNE, « Moving horizon observers and observer-based control », in : *IEEE Transactions on Automatic Control* 40.6 (1995), p. 995-1006, DOI : 10.1109/9.388677.
- [Mic+18] Paulius MICEKEVICIUS et al., « Mixed Precision Training », in : *International Conference on Learning Representations*, 2018, URL : <https://openreview.net/forum?id=r1gs9JgRZ>.
- [Mir+24] Iman MIRZADEH et al., « GSM-Symbolic : Understanding the Limitations of Mathematical Reasoning in Large Language Models », in : *arXiv preprint arXiv :2410.05229* (2024).
- [Mou88] Hervé MOULIN, « Axioms of Cooperative Decision Making », in : 1988.
- [MWZ22] Simon MUNTWILER, Kim P. WABERSICH et Melanie N. ZEILINGER, « Learning-based Moving Horizon Estimation through Differentiable Convex Optimization Layers », in : *Proceedings of The 4th Annual Learning for Dynamics and Control Conference*, sous la dir. de Roya FIROOZI et al., t. 168, Proceedings of Machine Learning Research, PMLR, 23–24 Jun 2022, p. 153-165.
- [Neg15] Elsa NEGRE, *Information and Recommender Systems*, John Wiley & Sons, 2015.
- [NET20] Andi NIKA, Sepehr ELAHI et Cem TEKIN, « Contextual Combinatorial Volatile Multi-armed Bandit with Adaptive Discretization », in : *The 23rd International Conference on Artificial Intelligence and Statistics, AISTATS 2020, 26-28 August 2020, Online [Palermo, Sicily, Italy]*, sous la dir. de Silvia CHIAPPA et Roberto CALANDRA, t. 108, Proceedings of Machine Learning Research, PMLR, 2020, p. 1486-1496.
- [Oco+01] Mark O'CONNOR et al., « PolyLens : a recommender system for groups of users », in : *ECSCW 2001*, Springer, 2001, p. 199-218.
- [Öne+18] Doruk ÖNER et al., « Combinatorial Multi-Objective Multi-Armed Bandit Problem », in : *CoRR* abs/1803.04039 (2018), arXiv : 1803.04039.
- [Oor+18] Sageev OORE et al., « This Time with Feeling : Learning Expressive Musical Performance », in : *Neural Computing and Applications* 32 (2018), p. 955-967, URL : <https://link.springer.com/article/10.1007/s00521-018-3758-9>.

-
- [Oze+25] Derin OZER et al., « A Transformer Model for Predicting Chemical Reaction Products from Generic Templates », in : *arXiv preprint arXiv :2503.05810* (2025).
- [Pac+06] H.B. PACEJKA et al., *Tyre and Vehicle Dynamics*, Elsevier Science, 2006.
- [PS13] Christos H PAPADIMITRIOU et Kenneth STEIGLITZ, *Combinatorial optimization : algorithms and complexity*, Courier Corporation, 2013.
- [Pay19] Christine PAYNE, *MuseNet*, 2019, URL : <https://openai.com/blog/musenet>.
- [PB07] Michael J. PAZZANI et Daniel BILLSUS, « Content-based recommendation systems », in : *The adaptive web*, Springer, 2007, p. 325-341.
- [PM18] Judea PEARL et Dana MACKENZIE, *The book of why : the new science of cause and effect*, Basic books, 2018.
- [Pei+23] Jiyuan PEI et al., « Local optima correlation assisted adaptive operator selection », in : *Proceedings of the Genetic and Evolutionary Computation Conference, 2023*, p. 339-347.
- [PLD05] Hanchuan PENG, Fuhui LONG et Chris DING, « Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy », in : *IEEE Transactions on pattern analysis and machine intelligence* 27.8 (2005), p. 1226-1238.
- [Per82] Alan J PERLIS, « Special feature : Epigrams on programming », in : *ACM Sigplan Notices* 17.9 (1982), p. 7-13.
- [Pic11] Romain PICOT-CLÉMENTE, « Une architecture générique de Systèmes de recommandation de combinaison d'items : application au domaine du tourisme », Theses, Université de Bourgogne, 2011, URL : <https://tel.archives-ouvertes.fr/tel-00688994>.
- [Pin99] Allan PINKUS, « Approximation theory of the MLP model in neural networks », in : *Acta Numerica* 8 (1999), p. 143-195.
- [Pop13] George POPESCU, « Group recommender systems as a voting problem », in : *International Conference on Online Communities and Social Computing*, Springer, 2013, p. 412-421.
- [Pru21] Geoffrey PRUVOST, « Contributions à l'optimisation multi-objectif à base de décomposition », Theses, Université de Lille, déc. 2021.
- [Rad+19] Alec RADFORD et al., « Language models are unsupervised multitask learners », in : *OpenAI blog* 1.8 (2019), p. 9.
- [RKJ08] Filip RADLINSKI, Robert KLEINBERG et Thorsten JOACHIMS, « Learning diverse rankings with multi-armed bandits », in : *Machine Learning, Proceedings of the Twenty-Fifth International Conference (ICML 2008), Helsinki, Finland, June 5-9, 2008*, sous la dir. de William W. COHEN, Andrew MCCALLUM et Sam T. ROWEIS, t. 307, ACM International Conference Proceeding Series, ACM, 2008, p. 784-791, DOI : 10.1145/1390156.1390255.

-
- [RPK19] Maziar RAISSI, Paris PERDIKARIS et George E KARNIADAKIS, « Physics-informed neural networks : A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations », in : *Journal of Computational physics* 378 (2019), p. 686-707.
- [Raj11] Rajesh RAJAMANI, *Vehicle dynamics and control*, Springer Science & Business Media, 2011.
- [RRM03] Christopher V RAO, James B RAWLINGS et David Q MAYNE, « Constrained state estimation for nonlinear discrete-time systems : Stability and moving horizon approximations », in : *IEEE transactions on automatic control* 48.2 (2003), p. 246-258.
- [RW06] Carl Edward RASMUSSEN et Christopher K. I. WILLIAMS, *Gaussian processes for machine learning*. Adaptive computation and machine learning, MIT Press, 2006, p. I-XVIII, 1-248, ISBN : 026218253X.
- [RG19] Nils REIMERS et Iryna GUREVYCH, « Sentence-BERT : Sentence Embeddings using Siamese BERT-Networks », in : *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*, Association for Computational Linguistics, nov. 2019, URL : <https://arxiv.org/abs/1908.10084>.
- [Ren+20] Yi REN et al., « PopMAG : Pop Music Accompaniment Generation », in : *Proceedings of the 28th ACM International Conference on Multimedia*, Seattle, WA, USA : Association for Computing Machinery, 2020, p. 1198-1206, ISBN : 9781450379885, DOI : 10.1145/3394171.3413721, URL : <https://doi.org/10.1145/3394171.3413721>.
- [Rev+22] Guy REVACH et al., « KalmanNet : Neural Network Aided Kalman Filtering for Partially Known Dynamics », in : *Trans. Sig. Proc.* 70 (jan. 2022), p. 1532-1547, ISSN : 1053-587X, DOI : 10.1109/TSP.2022.3158588.
- [RRS15] Francesco RICCI, Lior ROKACH et Bracha SHAPIRA, « Recommender systems : introduction and challenges », in : *Recommender systems handbook*, Springer, 2015, p. 1-34.
- [RRS21] Francesco RICCI, Lior ROKACH et Bracha SHAPIRA, « Recommender systems : Techniques, applications, and challenges », in : *Recommender systems handbook* (2021), p. 1-35.
- [Ric+76] John R. RICE et al., « The algorithm selection problem », in : *Advances in computers* 15.65-118 (1976), p. 5, DOI : 10.1016/S0065-2458(08)60520-3.
- [Roa96] Curtis ROADS, *Computer Music Tutorial*, t. 170, MIT Press, 1996.
- [Rob52] Herbert E. ROBBINS, « Some aspects of the sequential design of experiments », in : *Bulletin of the American Mathematical Society* 58 (1952), p. 527-535.

-
- [Rob+18] Adam ROBERTS et al., « A Hierarchical Latent Vector Model for Learning Long-Term Structure in Music », in : *Proceedings of the 35th International Conference on Machine Learning*, sous la dir. de Jennifer DY et Andreas KRAUSE, t. 80, Proceedings of Machine Learning Research, PMLR, oct. 2018, p. 4364-4373, URL : <https://proceedings.mlr.press/v80/roberts18a.html>.
- [RH10] David ROGERS et Mathew HAHN, « Extended-connectivity fingerprints », in : *Journal of chemical information and modeling* 50.5 (2010), p. 742-754.
- [RZN17] Diederik M. ROIJERS, Luisa M. ZINTGRAF et Ann NOWÉ, « Interactive Thompson Sampling for Multi-objective Multi-armed Bandits », in : *Algorithmic Decision Theory - 5th International Conference, ADT 2017, Luxembourg, Luxembourg, October 25-27, 2017, Proceedings*, sous la dir. de Jörg ROTHE, t. 10576, Lecture Notes in Computer Science, Springer, 2017, p. 18-34, DOI : 10.1007/978-3-319-67504-6_2.
- [Rud+17] Samuel H RUDY et al., « Data-driven discovery of partial differential equations », in : *Science advances* 3.4 (2017), e1602614.
- [Rüt+23] Dimitri von RÜTTE et al., « FIGARO : Controllable Music Generation using Learned and Expert Features », in : *The Eleventh International Conference on Learning Representations, 2023*, URL : <https://openreview.net/forum?id=NyR80ZFHw6i>.
- [Sch+23] Julian D. SCHILLER et al., « A Lyapunov Function for Robust Stability of Moving Horizon Estimation », in : *IEEE Transactions on Automatic Control* 68.12 (2023), p. 7466-7481, DOI : 10.1109/TAC.2023.3280344.
- [SL19] Johan SCHOUKENS et Lennart LJUNG, « Nonlinear system identification : A user-oriented road map », in : *IEEE Control Systems Magazine* 39.6 (2019), p. 28-99.
- [Sch+08] Marcio SCHWAAB et al., « Nonlinear parameter estimation through particle swarm optimization », in : *Chemical Engineering Science* 63.6 (2008), p. 1542-1552, ISSN : 0009-2509, DOI : <https://doi.org/10.1016/j.ces.2007.11.024>, URL : <https://www.sciencedirect.com/science/article/pii/S0009250907008755>.
- [STS09] Maria SECRIER, Tina TONI et Michael P. H. STUMPF, « The ABC of reverse engineering biological signalling systems », in : *Mol. BioSyst.* 5 (12 2009), p. 1925-1935, DOI : 10.1039/B908951A, URL : <http://dx.doi.org/10.1039/B908951A>.
- [Sed18] Florence SEDES, *How information systems can help in alarm/alert detection*, Elsevier, 2018.
- [SHB16] Rico SENNRICH, Barry HADDOW et Alexandra BIRCH, « Neural Machine Translation of Rare Words with Subword Units », in : *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1 : Long Papers)*, Berlin, Germany : Association for Computational Linguistics, août 2016, p. 1715-1725, DOI : 10.18653/v1/P16-1162, URL : <https://aclanthology.org/P16-1162>.

-
- [SUV18] Peter SHAW, Jakob USZKOREIT et Ashish VASWANI, « Self-attention with relative position representations », in : *arXiv preprint arXiv :1803.02155* (2018).
- [Shi+99] Yusuxke SHIBATA et al., « Byte pair encoding : A text compression scheme that accelerates pattern matching », in : (1999).
- [Sin89] Naresh K. SINHA, « System identification - Theory for the user : Lennart Ljung. », in : *Autom.* 25.3 (1989), p. 475-476, URL : <http://dblp.uni-trier.de/db/journals/automatica/automatica25.html#Sinha89>.
- [Smi09] Kate A SMITH-MILES, « Cross-disciplinary perspectives on meta-learning for algorithm selection », in : *ACM Computing Surveys (CSUR)* 41.1 (2009), p. 1-25, DOI : 10.1145/1456650.1456656.
- [Sto98] W. J. H. STORTELDER, *Parameter estimation in nonlinear dynamic systems*, eng, CWI tract ; 124, Amsterdam, Netherlands : Centrum voor Wiskunde en Informatica, 1998, ISBN : 9061964822.
- [SSK15] Bob L. STURM, João Felipe SANTOS et Iryna KORSHUNOVA, « Folk Music Style Modelling by Recurrent Neural Networks with Long Short-Term Memory Units », in : *Extended abstracts for the Late-Breaking Demo Session of the 16th International Society for Music Information Retrieval Conference*, 2015, URL : <https://ismir2015.ismir.net/LBD/LBD13.pdf>.
- [SB98] Richard S. SUTTON et Andrew G. BARTO, *Reinforcement learning - an introduction*, Adaptive computation and machine learning, MIT Press, 1998, ISBN : 978-0-262-19398-6.
- [Tay+21] Yi TAY et al., « Long Range Arena : A Benchmark for Efficient Transformers », in : *International Conference on Learning Representations*, 2021, URL : <https://openreview.net/forum?id=qVyeW-grC2k>.
- [Thi+23] John THICKSTUN et al., *Anticipatory Music Transformer*, 2023, arXiv : 2306.08620 [cs.SD].
- [Thi05] Dirk THIERENS, « An adaptive pursuit strategy for allocating operator probabilities », in : *Proceedings of the 7th annual conference on Genetic and evolutionary computation*, 2005, p. 1539-1546.
- [Tho33] William R. THOMPSON, « On the Likelihood that One Unknown Probability Exceeds Another in View of the Evidence of Two Samples », in : *Biometrika* 25.3/4 (1933), p. 285, ISSN : 0006-3444, DOI : 10.2307/2332286.
- [TFF24] Giacomo di TOLLO, Gerarda FATTORUSO et Gianni FILOGRASSO, « An adaptive evolutionary strategy for long-short portfolio construction », in : *Decisions in Economics and Finance* (2024), p. 1-22.
- [Ton+09] Tina TONI et al., « Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems », in : *Journal of the Royal Society Interface* 6.31 (2009), p. 187-202.

-
- [Tur09] Alan M TURING, *Computing machinery and intelligence*, Springer, 2009.
- [VPM20] Jyoti VASHISHTHA, Vijay Hasan PURI et MUKESH, « Feature Selection Using PSO : A Multi Objective Approach », in : *Machine Learning, Image Processing, Network Security and Data Sciences : Second International Conference, MIND 2020, Silchar, India, July 30-31, 2020, Proceedings, Part II 2*, Springer, 2020, p. 106-119.
- [Vas+17] Ashish VASWANI et al., « Attention is All you Need », in : *Advances in Neural Information Processing Systems*, sous la dir. d'I. GUYON et al., t. 30, Curran Associates, Inc., 2017, URL : <https://proceedings.neurips.cc/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf>.
- [WOJ17] Handing WANG, Markus OLHOFFER et Yaochu JIN, « A mini-review on preference modeling and articulation in multi-objective optimization : current status and challenges », in : *Complex Intelligent Systems (2017)*, DOI : 10.1007/s40747-017-0053-9.
- [WX06] Shengwei WANG et Xinhua XU, « Parameter estimation of internal thermal mass of building dynamic models using genetic algorithm », in : *Energy Conversion and Management* 47.13 (2006), p. 1927-1941, ISSN : 0196-8904, DOI : <https://doi.org/10.1016/j.enconman.2005.09.011>, URL : <https://www.sciencedirect.com/science/article/pii/S0196890405002311>.
- [Wan+20] Ziyu WANG et al., « POP909 : A Pop-song Dataset for Music Arrangement Generation », in : *Proceedings of 21st International Conference on Music Information Retrieval, ISMIR, 2020*, URL : <https://arxiv.org/abs/2008.07142>.
- [WYH23] Wenpeng WEI, Guodong YIN et Tianyi HE, « Physics-informed data-based lpv modeling and validations of lateral vehicle dynamics », in : *IEEE Transactions on Intelligent Vehicles (2023)*.
- [Wei17] W. WEINAN, « A Proposal on Machine Learning via Dynamical Systems », English (US), in : *Communications in Mathematics and Statistics* 5.1 (mars 2017), p. 1-11, ISSN : 2194-6701.
- [WKW16] Karl WEISS, Taghi M KHOSHGOFTAAR et DingDing WANG, « A survey of transfer learning », in : *Journal of Big data* 3 (2016), p. 1-40.
- [Wil+20] Jared WILLARD et al., « Integrating physics-based modeling with machine learning : A survey », in : *arXiv preprint arXiv :2003.04919 1.1* (2020), p. 1-34.
- [Wol+20] Thomas WOLF et al., « Transformers : State-of-the-Art Natural Language Processing », in : *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing : System Demonstrations*, Online : Association for Computational Linguistics, oct. 2020, p. 38-45, URL : <https://www.aclweb.org/anthology/2020.emnlp-demos.6>.

-
- [Wu+15] Yiqi WU et al., « Service-oriented feature-based data exchange for cloud-based design and manufacturing », in : *IEEE Transactions on services computing* 11.2 (2015), p. 341-353.
- [Wu+16] Yonghui WU et al., *Google's Neural Machine Translation System : Bridging the Gap between Human and Machine Translation*, 2016, arXiv : 1609.08144 [cs.CL].
- [YL20] Li-Chia YANG et Alexander LERCH, « On the Evaluation of Generative Models in Music », in : *Neural Comput. Appl.* 32.9 (mai 2020), p. 4773-4784, ISSN : 0941-0643, DOI : 10.1007/s00521-018-3849-7, URL : <https://doi.org/10.1007/s00521-018-3849-7>.
- [Yan+18] Xiaomin YANG et al., « Multiple dictionary pairs learning and sparse representation-based infrared image super-resolution with improved fuzzy clustering », in : *Soft Computing* 22 (2018), p. 1385-1398.
- [Yin23] Yuan YIN, « Physics-Aware Deep Learning and Dynamical Systems : Hybrid Modeling and Generalization », thèse de doct., Sorbonne Université, 2023.
- [Yon+19] Jia-shi YONG et al., « A novel bat algorithm based on cross boundary learning and uniform explosion strategy », in : *Applied Mathematics-A Journal of Chinese Universities* 34.4 (2019), p. 480-502.
- [YXW14] Yuan YUAN, Hua XU et Bo WANG, « An improved NSGA-III procedure for evolutionary many-objective optimization », in : *Proceedings of the 2014 annual conference on genetic and evolutionary computation*, 2014, p. 661-668.
- [ZB09] Victor M ZAVALA et Lorenz T BIEGLER, « Nonlinear programming strategies for state estimation and model predictive control », in : *Nonlinear Model Predictive Control : Towards New Challenging Applications* (2009), p. 419-432.
- [Zen+21] Mingliang ZENG et al., « MusicBERT : Symbolic Music Understanding with Large-Scale Pre-Training », in : *Findings of the Association for Computational Linguistics : ACL-IJCNLP 2021*, Online : Association for Computational Linguistics, août 2021, p. 791-800, DOI : 10.18653/v1/2021.findings-acl.70, URL : <https://aclanthology.org/2021.findings-acl.70>.
- [ZL07] Qingfu ZHANG et Hui LI, « MOEA/D : A multiobjective evolutionary algorithm based on decomposition », in : *IEEE Transactions on evolutionary computation* 11.6 (2007), p. 712-731.
- [Zha+16] Xingyi ZHANG et al., « A decision variable clustering-based evolutionary algorithm for large-scale many-objective optimization », in : *IEEE Transactions on evolutionary computation* 22.1 (2016), p. 97-112.
- [ZW22] Yong ZHENG et David (Xuejun) WANG, « A survey of recommender systems with multi-objective optimization », in : *Neurocomputing* 474 (2022), p. 141-153, ISSN : 0925-2312, DOI : <https://doi.org/10.1016/j.neucom.2021.11.041>, URL : <https://www.sciencedirect.com/science/article/pii/S0925231221017185>.

-
- [ZHO17] Hongliang ZHONG, « Bandit feedback in Classification and Multi-objective Optimization », Theses, Mathématique et Informatique de Marseille (ED184), 2017.
- [Zhu+19] Yinhao ZHU et al., « Physics-constrained deep learning for high-dimensional surrogate modeling and uncertainty quantification without labeled data », in : *Journal of Computational Physics* 394 (2019), p. 56-81.
- [Zit01] E ZITZLER, « SPEA2 : Improving the Strength Pareto Evolutionary Algorithm », in : (2001).
- [ZK04] Eckart ZITZLER et Simon KÜNZLI, « Indicator-based selection in multiobjective search », in : *International conference on parallel problem solving from nature*, Springer, 2004, p. 832-842.

Annexes

CONTRIBUTIONS COMPLÉMENTAIRES ET INFORMATIONS COMPLÉMENTAIRES SUR LES CONTRIBUTIONS.

A.1 Exploitation de retours utilisateur partiels

La seconde stratégie de nos contributions consistait à palier la parcimonie des données (retours partiels des utilisateurs) pour les rendre plus exploitables par les modèles d'apprentissage¹.

De 2020 à 2024, nous avons développé et amélioré des approches innovantes telles que *P-BSB* (*Partial Bandit with Semi-Bandit*) [Let+20b] et *BUSBC* (*Bandit Under Semi-Bandit Conditions*) [Let+22]. Ces méthodes visent à maximiser l'efficacité des systèmes de recommandation dans des environnements où les retours d'utilisateurs sont rares ou partiels. Elles s'appuient sur des techniques de bandits-manchots combinatoires pour optimiser la sélection des recommandations tout en minimisant le besoin de retours exhaustifs des utilisateurs : elles sont conçues pour être robustes face aux variations des données (retours utilisateurs, évaluation des recommandations) et à leur faible densité (parcimonie).

Le développement de ces approches repose sur deux constats principaux : 1) les systèmes de recommandation nécessitent souvent de fonctionner avec des données partielles ou incertaines, ce qui rend difficile l'optimisation de l'apprentissage en ligne ; 2) l'efficacité des recommandations dépend fortement de la capacité à équilibrer exploration et exploitation avec un feedback utilisateur limité [Let23]. En conséquence, *P-BSB* et *BUSBC* ont été conçues pour tirer le meilleur parti des retours utilisateurs limités tout en conservant une bonne performance globale des algorithmes de recommandation.

A.1.1 *BSB* et *BUSBC*

Sélection de retours utilisateur. Dans le cadre de nos travaux sur la sélection des retours utilisateur, nous avons examiné l'approche traditionnelle, modélisée par la stratégie "*Reinforce - RE*" (voir équation A.1). En parallèle, nous avons développé deux nouvelles méthodes : *Optimal-Exploration (OE)* (voir équation A.2) et *Randomized (RD)* (voir équation A.3) [Let+20a ; Let+20b].

1. Celle-ci a été expérimenté dans le cadre mono-objectif mais peut très bien s'appliquer au cadre bi-objectif.

Ci-dessous un récapitulatif de ces stratégies

- **Reinforce - RE** : sélectionne les ψ bras de S_t ayant l'espérance de récompense $\mathbb{E}[R_{t,a}|x_t]$ la plus haute, c.-à-d. :

$$a_i = \operatorname{argmax}_{a \in S_t \setminus \{P_t^{-1}\}} \mathbb{E}[R_{t,a}|x_t] \quad (\text{A.1})$$

- **Optimal-Exploration - OE** : sélectionne les ψ bras de S_t ayant été le moins fréquemment observés à l'itération t , c.-à-d. :

$$a_i = \operatorname{argmin}_{a \in S_t \setminus \{P_t^{-1}\}} \text{obs}_{a,t} \quad (\text{A.2})$$

où $\text{obs}_{a,t}$ est le nombre de fois qu'un retour utilisateur a été émis pour le bras a à l'itération t .

- **Randomized - RD** : sélectionne aléatoirement ψ bras distincts de S_t , c.-à-d. :

$$a_i = \operatorname{random}(S_t \setminus \{P_t^{-1}\}) \quad (\text{A.3})$$

où $\operatorname{random}(S_t \setminus \{P_t^{-1}\})$ correspond à la sélection aléatoire d'un bras dans S_t n'ayant pas été précédemment choisi.

Ces stratégies peuvent aussi s'appliquer aux récompenses différées [HL16].

Méthodes BSB et BUSBC. Nous avons visé des méthodes applicables à un large éventail d'applications, sans dépendance à des prérequis spécifiques. Seuls les identifiants des éléments de \mathcal{A} et les opinions des utilisateurs sont considérés. Nos approches peuvent donc être intégrées dans tout système de recommandation utilisant un algorithme de bandit-manchot.

Inspirés par la littérature sur les bandits-manchots combinatoires et les retours utilisateurs de type bandit et semi-bandit [ABL11], nous avons proposé deux nouvelles méthodes :

- **Bandit and Semi-Bandit (BSB)** [Let+20a ; Let+20b] : combine l'observation de la récompense cumulée $R_t^B = P_t^\top F_t$ pour chaque bras $a_i \in S_t$ et le retour utilisateur $F_{t,a}$. Cela permet une double récompense pour les éléments jugés pertinents.
- **Bandit under Semi-Bandit Conditions (BUSBC)** [Let+21 ; Let+22] : n'observe la récompense cumulée R_t^B que si le retour utilisateur F_{t,a_i} , spécifique au bras a_i , indique un intérêt. Conçue pour les algorithmes à borne supérieure de confiance, cette méthode restreint l'attribution des récompenses pour améliorer la convergence.

Les deux méthodes visent à réduire les temps de convergence par une manipulation des récompenses observées. Notamment, *BUSBC* est efficace pour les algorithmes optimistes tels que *UCB1*, où la récompense cumulée n'est attribuée qu'aux bras confirmés comme pertinents par l'utilisateur.

L'approche *BSB* attribue une récompense cumulée à tous les bras de S_t et, si l'opinion de l'utilisateur est observée, une récompense individuelle supplémentaire est ajoutée :

$$R_t^B = P_t^\top F_t, \quad \forall a \in P_t : SR_{t,a} = SR_{t-1,a} + R_t^B + F_{t,a} \quad (\text{A.4})$$

Cependant, cette méthode peut entraîner une attribution excessive de récompenses, limitant la capacité de certains algorithmes à ajuster dynamiquement la composition des recommandations. Par exemple, avec *UCB1* dans certains cas (p.ex., 18 bras, $k = 6$), tous les éléments d'un super-bras S_t pourraient recevoir la même récompense, compliquant l'identification des meilleures combinaisons.

Pour pallier ce problème, *BUSBC* limite l'attribution des récompenses cumulées aux seuls bras pertinents, favorisant un apprentissage basé sur des observations réelles :

$$\forall a \in P_t, \text{ si } F_{a,t} = 1 : SR_{a,t} = SR_{a,t-1} + R_t^B \quad (\text{A.5})$$

Cette méthode, optimisée pour les approches à borne de confiance, augmente la vitesse de convergence proportionnellement à la pertinence globale de S_t .

Exemple illustratif : Pour S_t de $k = 8$ éléments avec $F_t = \{0, 1, 0, 1\}$, les récompenses calculées par *BSB* et *BUSBC* diffèrent comme suit :

- **BSB** : tous les bras reçoivent $R_t^B = 2$. Les éléments observés reçoivent des récompenses supplémentaires ;
- **BUSBC** : seuls les bras confirmés pertinents reçoivent $R_t^B = 2$, les autres restent inchangés.

Ainsi, *BSB* suit une approche optimiste, attribuant des récompenses même sans retour direct, tandis que *BUSBC* est plus réaliste, se basant uniquement sur des observations vérifiées.

Une illustration du fonctionnement de l'algorithme *BUSBC*, avec l'exemple de *PV-BUSBC-OE*, est fournit à l'annexe B.7.

A.1.2 Synthèse des résultats obtenus

Les résultats montrent que chaque étape du modèle de stratégie de feedback proposé (Identification du Feedback, Récupération du Feedback, et Calcul de la Récompense) peut être optimisée pour améliorer les performances des algorithmes *COM-MAB*. En particulier, le processus de calcul de la récompense via *BUSBC* apporte des gains notables pour les algorithmes à borne supérieure de confiance (*UCB*) en améliorant la précision globale et en accélérant la convergence.

Contrairement à *BSB*, *BUSBC* attribue les récompenses cumulées uniquement aux bras jugés pertinents, évitant ainsi les récompenses injustifiées et renforçant l'efficacité de l'apprentissage. Bien que conçu pour les algorithmes de type *UCB*, *BUSBC* demeurent adaptables et performants avec d'autres méthodes. Ainsi, ils s'avèrent être le choix le plus approprié pour la plupart des scénarios, en particulier pour les approches optimistes comme *UCB*.

A.1.3 Synthèse et analyse réflexive sur les méthodes *P-BSB* et *BUSBC*

Les approches *P-BSB* et *BUSBC* incluent plusieurs variantes adaptées aux différents scénarios de recommandation et objectifs d'optimisation. Parmi ces méthodes, on retrouve :

- *P-BSB (Partial Bandit with Semi-Bandit)* [Let+20b] - Cette méthode combine les stratégies de bandit et de semi-bandit en sélectionnant dynamiquement un sous-ensemble d'éléments pour lesquels des retours utilisateurs seront sollicités. Elle permet de réduire le nombre de sollicitations nécessaires tout en maintenant une précision acceptable grâce à une approche hybride de retour d'information. Cette méthode est particulièrement efficace dans des environnements où les données sont fragmentées ou coûteuses à acquérir.
- *BUSBC (Bandit Under Semi-Bandit Conditions)* - Cette approche propose un cadre de calcul de récompense adapté aux conditions de retour partiel d'information (semi-bandit). En ajustant la manière dont les récompenses sont agrégées, *BUSBC* permet d'améliorer la précision globale des algorithmes de recommandation basés sur les bandits combinatoires (*COM-MAB*) en contexte semi-bandit. Les expériences menées montrent que cette méthode surpasse les approches classiques de bandits-manchots, en particulier lorsque les informations disponibles sont partielles, augmentant la précision de manière significative (jusqu'à 16,2%). Cette méthode est particulièrement utile pour les recommandations en ligne nécessitant une gestion robuste de l'incertitude et de la non-stationnarité.

Pour ces deux méthodes de calcul de récompense (*BSB* et *BUSBC*) nous avons proposés les stratégies de sélection des retours utilisateurs suivantes :

- *Reinforce* pour sélectionner les éléments avec les plus fortes attentes de récompense ;
- *Optimal Exploration* pour maximiser l'apprentissage des récompenses des éléments moins observés ;
- *Randomized* pour un équilibre entre exploration et exploitation.

En conclusion, ces méthodes visent à améliorer la performance des systèmes de recommandation dans des environnements avec des données partielles ou incertaines, tout en minimisant le besoin de retours utilisateurs exhaustifs. Elles permettent d'équilibrer exploration et exploitation de manière flexible et d'optimiser plusieurs critères simultanément (précision, diversité) dans des contextes d'application variés. Ces approches offrent des garanties empiriques robustes pour les applications de recommandation en ligne, où la capacité d'adaptation et l'efficacité du calcul des récompenses sont primordiales.

A.2 Exemple de *Budgeted-Gorthaur* avec *EXP3-754515*.

L'algorithme *Budgeted-Gorthaur* décrit un cadre généraliste. Cependant, une implémentation possible de *Budgeted-Gorthaur* est l'*EXP3-754515* comme nous l'avons proposé dans [Let+24], c.-à-d., trois instances d'*EXP3* sont utilisées comme méta-algorithmes. Chaque méta-algorithme considère un vecteur de référence différent caractérisé par des angles Θ_i de : 75° ($\frac{5\pi}{12}$) : favori-

sant la précision globale), 45° ($\frac{\pi}{4}$: favorisant la performance équilibrée), et 15° ($\frac{\pi}{12}$: favorisant la diversité agrégée) respectivement.

Chacun de ces méta-algorithmes sélectionne indépendamment un algorithme du portefeuille pour identifier k'_i éléments à recommander.

Il est possible de considérer les budgets k'_i comme des données d'entrée directement. Cependant, dans cette étude, les budgets sont déterminés dynamiquement en fonction des valeurs de k et d'un facteur de pondération W_i , où $\forall W_i \in W, W_i \in [0, 1]$ et $\sum_{i=1}^{|W|} W_i = 1$, ce qui indique la part de S_t attribuée à l'algorithme sélectionné par le méta-algorithme favorisant le compromis i . Le calcul de k'_i est donné par l'Équation A.6 ci-dessous, qui garantit que chaque compromis entre les critères est pris en compte et que la distribution d'entrée est au mieux approchée :

$$k'_i = \begin{cases} 1 & \text{si } k = |\Theta| \\ k - \sum_{j=0}^{i-1} k'_j & \text{si } i = |\Theta| \\ \lfloor kW_i \rfloor & \text{si } k \leq 2 |\Theta| \\ \lceil kW_i \rceil & \text{sinon} \end{cases} \quad (\text{A.6})$$

A.3 Compléments d'information sur l'algorithme génétique utilisé dans la méthode GAWC

Précisions concernant les détails spécifiques de configuration de l'algorithme génétique 4 :

- **Entrée** : n , le nombre d'individus dans chaque génération est fixé à 200, la taille d'un individu *individualSize* dépend du problème, le nombre de générations, *nbGens*, est fixé à 100, enfin, *mutationRate* est fixé à 0.10.
- **Sortie** : Après l'exécution complète, l'algorithme renvoie l'individu avec la meilleure fitness rencontré au cours de toutes les générations.
- **Initialisation** : Chaque individu de la première génération, un vecteur de taille *individualSize*, est généré aléatoirement.
- **Sélection par tournoi** : la population est regroupée aléatoirement par paires. La fitness des deux parents de chaque paire est comparée et seul le parent avec la meilleure fitness est conservé.
- **Croisement** : la population est regroupée aléatoirement en paires. Une opération de croisement est effectuée entre les deux parents de chaque paire pour obtenir la génération des enfants. Un croisement à deux points est utilisé : deux points de croisement sont choisis aléatoirement chez les deux parents. Les bits sont échangés entre les points de croisement pour obtenir les deux nouveaux enfants.

A.4 Quatre points de vue du Compass - Exemple concret

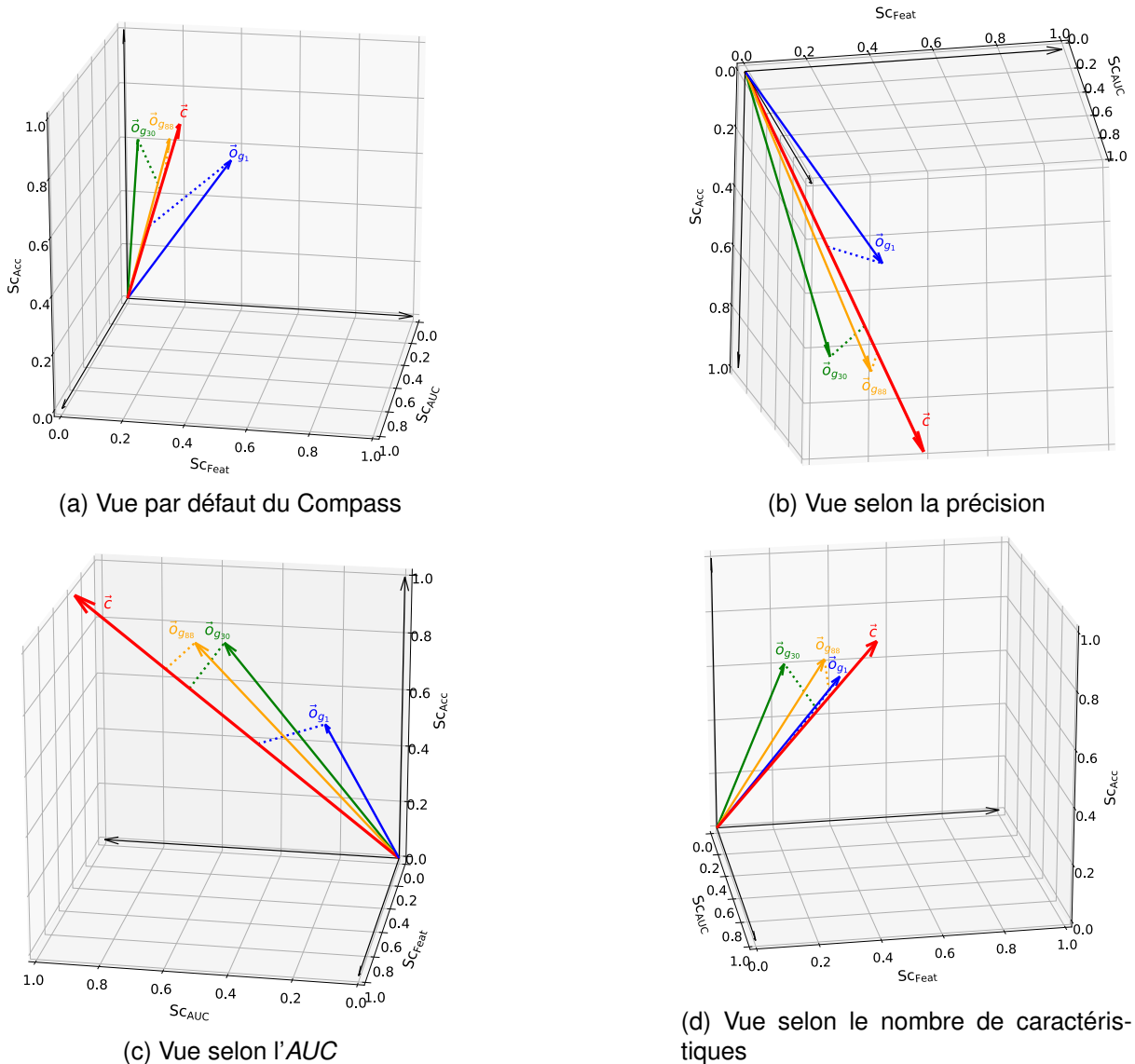


FIGURE A.1 – Résultats obtenus par le Compass pour 3 générations améliorées de g_1 à g_{88} (exemple d'une exécution réelle utilisant *ELM* comme classifieur sur le jeu de données Musk1), observables à travers quatre points de vue.

A.5 Tokenisation dans MidiTok

La Table A.1 montre une analyse comparative de différentes méthodes de tokenisation. Dans ces travaux originaux, les auteurs vantent les avantages d'utiliser ce que nous appelons des

tokens supplémentaires, qui représentent des informations autres que les notes et le temps. De plus, les tokenisations multi-vocabulaire ont été introduites avec pour objectif principal de réduire la longueur des séquences d’embeddings traitées par le modèle, ce qui peut devenir un goulot d’étranglement avec les modèles *Transformer*, dont la complexité augmente de manière quadratique.

Tokenisation	Temps		Durée des notes		Multi-piste	Flux unique	Multi-voc.	Accord	Silence	Tempo	Signature temp.	Pédale	Bend pitch
	TimeShift	Bar + Pos.	Duration	NoteOff									
()2-3 ()4-5													
<i>MIDI-Like</i> [Oor+18]	✓	-	-	✓	†	†	-	‡	‡	‡	‡	‡	‡
<i>REMI</i> [HY20]	-	✓	✓	-	†	†	-	‡	‡	‡	‡	‡	‡
<i>Structured</i> [HC21]	✓	-	-	-	†	†	-	-	-	-	-	-	-
<i>TSD</i> [Fra+23a]	✓	-	✓	-	†	†	-	‡	‡	‡	‡	‡	‡
<i>CP Word</i> [Hsi+21]	-	✓	✓	-	†	†	-	‡	‡	‡	‡	-	-
<i>Octuple</i> [Zen+21]	-	✓	✓	-	✓	✓	✓	-	-	‡	‡	-	-
<i>MuMIDI</i> [Ren+20]	-	✓	✓	-	✓	✓	✓	‡	-	‡	‡	-	-
<i>MMM</i> [EP20]	✓	✓	-	✓	✓	✓	-	‡	-	‡	‡	-	-

TABLE A.1 – Tableau comparatif des tokenisations implémentées par *MidiTok*. † : est vrai lorsque le tokenizer est configuré pour représenter les tokens *Program*; ‡ : est optionnel. Dans *MidiTok*, *MMM* est implémenté en utilisant des tokens *Duration*.

A.6 Tokens supplémentaires

MidiTok permet de choisir un ensemble de tokens supplémentaires à utiliser. Ces tokens ajoutent plus d’informations musicales, ce qui peut être utile dans certains cas pour modéliser :

- Accord : token de piste décrivant l’accord formé par les notes suivantes. Ce type de token peut aider à modéliser explicitement l’harmonie formée par les accords ;
- Programmes : token de piste informant du programme, c’est-à-dire l’instrument, des notes suivantes. Ce token est utilisé nativement dans certaines tokenisations ;
- Pédale de sustain : token de piste représentant les événements de pédale de sustain ;
- Bend de pitch : token de piste représentant les événements de bend de pitch ;
- Tempo : token global informant du tempo actuel, indiquant la vitesse d’exécution ;
- Signature rythmique : token global informant de la signature rythmique. La valeur de la signature rythmique impacte directement le nombre de temps présents dans les mesures, ainsi que la durée des temps ;
- Silence : token temporel agissant comme des décalages temporels, représentant les silences lorsqu’aucune note n’est actuellement jouée. Le silence est une information en soi utile aux musiciens. Le représenter explicitement pour un modèle pourrait probablement l’aider dans les tâches de modélisation musicale, bien qu’aucune recherche actuelle ne soutienne cette affirmation.

[HY20] rapportent que l’utilisation des tokens *Tempo* et *Chord* pour un *Transformer* génératif a produit des résultats de meilleure qualité, préférés par les évaluateurs humains.

PSEUDO-CODES D'ALGORITHMES DE L'ÉTAT DE L'ART EMPLOYÉS DANS NOS TRAVAUX

B.1 Algorithme ε -Greedy

Algorithme 8 : ε -Greedy pour la recommandation

Données : La liste des utilisateurs $u \in U$, la liste des éléments à recommander (c.-à-d., les bras) $a \in A$, le paramètre de sélection aléatoire $\varepsilon \in [0, 1]$.

```

1 pour  $t = 1$  à  $T$  faire
2   Sélectionner aléatoirement un utilisateur  $u \in U$ ;
3   si un élément  $a$  n'a pas encore été sélectionné alors
4     Recommander cet élément  $a$  à l'utilisateur  $u$ ;
5   sinon
6     Générer un nombre réel  $n$  entre 0 et 1;
7     si  $\varepsilon < n$  alors
8       Calculer  $\hat{\mu}_{a,t}$ ;
9       Recommander l'élément  $a_t = \arg \max (\hat{\mu}_{a,t})$  à l'utilisateur  $u$ ;
10    sinon
11      Recommander aléatoirement un élément  $a \in A$  à l'utilisateur  $u$ ;
12  Observer la récompense obtenue  $r_t$  et mettre à jour la moyenne  $\mu_{a,t}$ ;

```

B.2 Algorithme *UCB*

Algorithme 9 : *UCB1* pour la recommandation

Données : La liste des utilisateurs $u \in U$, la liste des éléments à recommander (c.-à-d., les bras) $a \in A$, le paramètre α ($\alpha=2$).

```
1 pour  $t = 1$  à  $T$  faire
2   Sélectionner aléatoirement un utilisateur  $u \in U$ ;
3   si un élément  $a \in A$  n'a pas encore été sélectionné alors
4     | Recommander cet élément  $a$  à l'utilisateur  $u$ ;
5   sinon
6     | Calculer pour chaque bras  $\hat{\mu}_{a,t}$  et la borne de confiance supérieure  $\sqrt{\frac{\alpha \ln t}{t_a}}$ ;
7     | Sélectionner l'élément  $a_t = \arg \max (\hat{\mu}_{a,t} + \sqrt{\frac{\alpha \ln t}{t_a}})$  et le recommander à
      | l'utilisateur  $u$ ;
8   Observer la récompense obtenue  $r_t$  et mettre à jour la moyenne  $\mu_{a,t}$ ;
```

B.3 Algorithme *Thompson Sampling*

Algorithme 10 : *Thompson Sampling (TS)* (Bernoulli) pour la recommandation

Données : La liste des utilisateurs $u \in U$, la liste des éléments à recommander (c.-à-d., les bras) $a \in A$, l'initialisation des succès et échecs $\forall a \in A, F_a = 0$ et $S_a = 0$.

```
1 pour  $t = 1$  à  $T$  faire
2   Sélectionner aléatoirement un utilisateur  $u \in U$ ;
3   pour tous les  $a \in A$  faire
4     | Générer l'échantillon  $\gamma_{t,a}$  à partir de la distribution :
      |
      | 
$$\text{Beta}(S_{t,a} + 1, F_{t,a} + 1)$$

      |
5   Sélectionner l'élément  $a_t = \arg \max (\gamma_{a,t})$  et le recommander à l'utilisateur  $u$ ;
6   Observer la récompense obtenue  $r_t$ ;
7   si  $r_t = 1$  alors
8     |  $S_a = S_a + 1$ 
9   sinon
10  |  $F_a = F_a + 1$ 
```

B.4 Algorithme *EXP3*

Algorithme 11 : *EXP3* pour la recommandation

Données : La liste des utilisateurs $u \in U$, la liste des k éléments à recommander $a \in A$ (c.-à-d., les bras), $\eta \in [0, 1]$, $\forall a \in A, w_a = 1$.

- 1 **pour** $t = 1$ à T **faire**
- 2 Sélectionner aléatoirement un utilisateur $u \in U$;
- 3 Recommander l'élément a_t à l'utilisateur u selon la probabilité :

$$P_{a,t} = (1 - \eta) \frac{w_{a,t}}{\sum_{i=1}^k w_{a_i,t}} + \frac{\eta}{k}$$

- 4 Observer la récompense $r_{a,t}$ retournée par l'utilisateur;
 - 5 **pour tous les** $a \in A$ **faire**
 - 6 $\hat{r}_{a,t} = \begin{cases} r_{a,t}/p_{a,t} & \text{si } a = a_t; \\ 0 & \text{sinon} \end{cases}$;
 - 7 Mettre à jour $w_{a,t+1} = w_{a,t} \exp\left(\eta \frac{\hat{r}_{a,t}}{k}\right)$;
-

B.5 Algorithme *LinUCB*

Algorithme 12 : *LinUCB* (*Linear Disjoint*)

Données : L'ensemble des k éléments à recommander (c.-à-d., les bras) $a \in A$ disponibles, $\alpha \in \mathbb{R}^+$, l'horizon T , et l'ensemble des n contextes fixes disponibles X . $\forall a \in A, M_a \leftarrow I_d, b_a \leftarrow 0_d$.

- 1 **pour** $t = 1$ à T **faire**
 - 2 Considérer $x_t \in X$: un utilisateur et son contexte;
 - 3 **pour tous les** $a \in A$ **faire**
 - 4 $\hat{\theta}_a \leftarrow M_a^{-1} b_a$;
 - 5 $p_{t,a} \leftarrow \hat{\theta}_a^\top x_t + \alpha \sqrt{x_t^\top M_a^{-1} x_t}$;
 - 6 Sélectionner l'élément $a_t = \arg \max_{a \in A} (p_{t,a})$ et le recommander à l'utilisateur u ;
 - 7 Observer la récompense r_t retournée par l'utilisateur u ;
 - 8 $M_{a_t} \leftarrow M_{a_t} + x_t x_t^\top$;
 - 9 $b_{a_t} \leftarrow b_{a_t} + r_t x_t$;
-

B.6 Algorithme *Linear Thompson Sampling*

Algorithme 13 : *Linear Thompson Sampling (LinTS)*

Données : L'ensemble des k éléments à recommander (c.-à-d., les bras) $a \in A$ disponibles, l'horizon T , et l'ensemble des n contextes fixes disponibles
 $X. \forall a \in A, B_a \leftarrow I_d, f_a \leftarrow 0_d, \hat{\theta}_a \leftarrow 0_d.$

```

1 pour  $t = 1$  à  $T$  faire
2   Considérer  $x_t \in X$  : un utilisateur et son contexte;
3   pour tous les  $a \in A$  faire
4     Générer l'échantillon  $\tilde{\theta}_{t,a}$  à partir de la distribution :
           
$$\mathcal{N}(\hat{\theta}_{t,a}, v^2 B_{t,a}^{-1})$$

5     Sélectionner l'élément  $a_t = \arg \max x_{t,a}^\top \tilde{\theta}_{t,a}$  et le recommander à l'utilisateur  $u$ ;
6     Observer la récompense  $r_t$  retournée par l'utilisateur  $u$ ;
7      $B_{t,a} \leftarrow B_{t,a} + x_{a,t} x_{a,t}^\top$ ;
8      $f_{t,a} \leftarrow f_{t,a} + r_t x_{a,t}$ ;
9      $\hat{\theta}_{t,a} \leftarrow B_{t,a}^{-1} f_{t,a}$ ;

```

B.7 Algorithme *PV-BUSBC-OE*

Algorithme 14 : *PV-BUSBC-OE*

Entrées : S_t : Super-bras recommandé.
 Y_t : Retours associés à \mathcal{A} ou utilisateur u_t .
 π : Politique de l'agent.
 ψ : Nombre de retours utilisateur attendus.

```

1  $P_t \leftarrow \emptyset$ 
2 tant que  $|P_t| < \psi$  faire
3   Construire  $P_t$  avec :
      $P_t = P_t \cup \{\operatorname{argmin}_{a \in S_t \setminus P_t} \operatorname{obs}_{a,t}\}$ 
4 fin
5 Acquérir les retours utilisateur :  $F_t = \{Y_{t,a} \mid a \in P_t\}$ 
6 Calculer la récompense cumulée :  $R_t^B = P_t^\top F_t$ 
7 pour  $a \in P_t$  faire
8   si  $F_{t,a} = 1$  alors
9     Mettre à jour la politique  $\pi$  avec :  $SR_{t,a} = SR_{t-1,a} + R_t^B$ 
10  fin
11 fin

```

LIENS DES PAGES WEB ARCHIVÉES

TABLE C.1 – Liens des pages archivées des sites web cités dans le mémoire¹

Page	Lien de la page archivée
Page principale	https://web.archive.org/web/https://github.com/ngutowski/
Repository algossim	https://web.archive.org/web/https://github.com/ngutowski/algossim
Repository GAWC	https://web.archive.org/web/https://github.com/ngutowski/gawc
Repository MidiTok	https://web.archive.org/web/https://github.com/Natooz/MidiTok
Repository article BPE	https://web.archive.org/web/https://github.com/Natooz/BPE-Symbolic-Music
Page personnelle	https://web.archive.org/web/https://ngutowski.fr
Site web de l'ESAIP	https://web.archive.org/web/https://www.esaip.org/
Site web de l'ESEO	https://web.archive.org/web/https://eseo.fr/
Site web de DIVA	https://web.archive.org/web/https://www.edih-diva.eu/
Séminaire IA & Musique	https://web.archive.org/web/https://leria.univ-angers.fr/?page_id=306

1. Hors liens directs vers les articles ou travaux publiés, hors lien vers mon profil *Google Scholar*

Titre : Apprentissage automatique pour la génération et l'exploration : Adaptation et contextualisation pour données complexes

Mot clés : Apprentissage par renforcement, Bandits-manchots, Optimisation multicritère, Évolutionnaire, Apprentissage profond, IA Générative

Resumé : Les travaux de recherche présentés dans ce mémoire se concentrent sur l'apprentissage automatique, avec une attention particulière portée sur l'adaptation et la contextualisation pour données complexes. La première partie traite de l'optimisation multicritère, avec une application des algorithmes de bandits-manchots aux systèmes de recommandation, et de méthodes évolutionnaires à la sélection d'attributs médicaux. La seconde partie se concentre sur les méthodes d'apprentissage profond : 1) pour l'estimation de paramètres ou d'état dans des systèmes dynamiques non linéaires ; 2) pour la génération appliquée à la musique symbolique. Enfin, la troisième partie présente les perspectives de recherche, notamment autour du projet ^GGENOCOD. Ce projet vise à tirer parti des compétences développées dans les travaux précédents pour la génération de nouveaux composés moléculaires.

Title : Machine Learning for generation and exploration : Adaptation and contextualization for complex data.

Keywords : Reinforcement learning, Multi-Armed Bandit, Multicriteria optimization, Evolutionary, Deep Learning, Generative AI

Abstract : The research presented in this habilitation dissertation focuses on machine learning, with particular attention to the adaptation and contextualization for complex data. The first part focuses on multicriteria optimization, with an application of multi-armed bandit algorithms to recommendation systems and evolutionary methods to medical feature selection. The second part focuses on deep learning methods : 1) for parameter ou state estimation in nonlinear dynamic systems ; 2) for generation applied to symbolic music. Finally, the third part outlines research perspectives, particularly focusing on the ^GGENOCOD project. This project aims to leverage the skills developed in previous work for the generation of new molecular compounds.