Apprentissage de modèles de préférence

Un mariage entre théorie de la décision et apprentissage automatique

Margot Hérin
Instituto de Telecomunicações, Lisbonne, Portugal

L'individu, le groupe ou l'entité en charge de *prendre une décision* se heurte souvent à une première difficulté : comparer des solutions réalisables qui, selon le point de vue considéré, présentent toutes des avantages et des inconvénients. Par exemple, la comparaison de propositions de projets de développement ou de politiques publiques (e.g., politiques climatiques [6], planification énergétique durable [17]) nécessite de tenir compte d'impacts de natures diverses (environnementaux, économiques, sociaux, etc.) ou d'avis, potentiellement divergents, de différents experts et parties prenantes. Un autre exemple, à l'échelle individuelle, est le choix d'un itinéraire de voyage, qui implique souvent de trouver un équilibre entre plusieurs critères conflictuels, tels que le coût, la durée, les émissions de carbone et le confort.

Cette thèse[1], ancrée dans le domaine de l'*intelligence artificielle*, s'intéresse à la construction de systèmes capables de faciliter ou d'automatiser l'arbitrage entre ces différents points de vue selon les *préférences* du décideur. En particulier, exploitant d'un côté des *modèles de préférence* issus de la *théorie de la décision*, et de l'autre des techniques d'*apprentissage automatique*, elle fournit des *algorithmes d'apprentissage de préférences*. Ces derniers

1. Réalisée au LIP6 (UMR 7606, CNRS), sous la direction de Patrice Perny (Sorbonne université, LIP6) et Nataliya Sokolovska (Sorbonne université, LCQB) et disponible en lecture complète ici : `https://theses.fr/2025SORUS169`.

DOI:10.48556/SIF.1024.27.131

permettent d'obtenir, à partir de l'observation d'exemples, des représentations formelles des préférences, rendant alors possible la comparaison et l'évaluation systématique des solutions, et, *in fine*, la prise de décision (par exemple, le choix de la meilleure solution).

Modélisation des préférences

En présence de n points de vue sur les solutions, on peut représenter celles-ci par des vecteurs $x = (x_1, \ldots, x_n)$, pris dans un ensemble admissible $\mathcal{X}$, représentant leurs conséquences selon chaque points de vue. Ensuite, une manière, parmi d'autres, de modéliser les préférences d'un décideur sur $\mathcal{X}$, peut être de faire appel à une *fonction d'utilité* $U_w : \mathcal{X} \to \mathbb{R}$, qui associe à chaque solution un score global. Le système de valeur du décideur guide alors cette évaluation à travers le paramètre w, encodant par exemple l'importance qu'il porte aux différents points de vue. Ainsi, après l'adaptation de w aux valeurs du décideur, une recommandation peut lui être proposée par l'optimisation de la fonction U_w sur l'ensemble $\mathcal{X}$ des solutions.

Ainsi, un vaste corpus de travaux dans le domaine de la *théorie de la décision*, allant de l'économie [20] et de la psychologie mathématique [19] à l'intelligence artificielle [1], s'est intéressé à la conception de ces fonctions d'utilité. L'objectif poursuivi est souvent double : assurer un *fort pouvoir descriptif* permettant de modéliser des comportements décisionnels variés et intégrant des logiques d'évaluation complexes, et garantir des propriétés mathématiques permettant d'assurer un certain niveau de *rationalité* et de cohérence parmi les décisions prises.

Une première approche élémentaire consiste à construire une *utilité additive* de la forme $\sum_{i=1}^{n} w_i u_i(x_i)$, où chaque conséquence x_i est d'abord encodée sur une échelle de satisfaction commune $[0, 1]$ par une *utilité marginale* u_i, puis pondérée par un poids w_i reflétant l'importance accordée au point de vue i. Cependant, l'agrégation des utilités marginales au moyen d'une somme pondérée limite fortement les capacités descriptives du modèle. Par exemple, cette agrégation ne permet pas de modéliser la préférence pour des solutions présentant des vecteurs d'utilité marginales *équilibrés*. Ainsi, il n'est pas possible de rendre $(0.5, 0.5)$ préférable simultanément à $(1, 0)$ et $(0, 1)$. Une solution simple dans ce cas consiste à attribuer un bonus aux solutions qui obtiennent de bonnes performances selon l'ensemble des points de vue. Il suffit pour cela d'ajouter un terme *conjonctif* à la somme pondérée, par exemple $\prod_{i=1}^{n} u_i(x_i)$ ou $\min\{u_1(x_1), \ldots, u_n(x_n)\}$, pondéré par un coefficient positif suffisamment élevé. On observe en effet que ce terme conjonctif prend une valeur strictement positive en $(0.5, 0.5)$, mais reste nul en $(1, 0)$ et en $(0, 1)$.

Plus généralement, au-delà de la partie linéaire de la fonction d'agrégation, divers termes d'*interaction* peuvent être introduits afin de modéliser des synergies positives, mais aussi négatives, au sein de certains groupes de points de vue. C'est l'idée des agrégateurs $F_w : [0,1]^n \mapsto \mathbb{R}$ paramétrés par une *fonction d'ensemble* w_S, $S \subseteq \{1,\ldots,n\}$ associant un poids à chaque groupe de critères potentiellement en interaction. De telles fonctions d'aggrégation permettent d'étendre le modèle de l'utilité additive à la classe des fonctions dites *totalement décomposables* (TD), s'écrivant sous la forme :

$$U_w(x) = F_w\left(u_1\left(x_1\right), \ldots, u_n\left(x_n\right)\right) \tag{1}$$

Deux exemples d'importance majeure de fonctions F_w sont l'*intégrale de Choquet* [8] et le modèle *multilinéaire* [16], qui modélisent respectivement les termes d'interaction à l'aide des opérateurs de conjonction $\min$ et $\prod$. Ces derniers agrégateurs sont croissants par rapport à chacun de leurs arguments lorsque w est monotone par rapport à l'inclusion des ensembles (on parle alors de *capacité*). Ceci permet de garantir qu'une certaine solution dont les performances sont supérieures à tout point de vue à celles d'une autre obtienne un score global plus élevé, i.e.

$$\forall i,\ u_i(x_i) \geq u_i(x'_i) \Rightarrow U_w(x) \geq U_w(x'). \tag{2}$$

Une autre classe importante de fonctions d'utilité autorisant les interactions entre points de vue et généralisant le modèle TD est celle des fonctions d'utilité GAI-décomposables [7]. Celles-ci prennent la forme d'une décomposition additive en termes multivariés :

$$U(x) = \sum_{S \in \mathcal{F}} u_S(x_S) \tag{3}$$

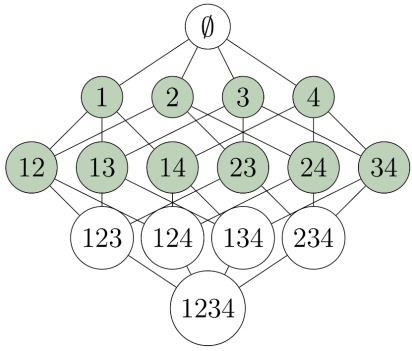
où $\mathcal{F}$ est une collection de sous-ensembles de $\{1,\ldots,n\}$, et où, pour tout $S \in \mathcal{F}$, x_S correspond au vecteur x restreint aux points de vue inclus dans S.

Apprentissage de représentations parcimonieuses de préférences

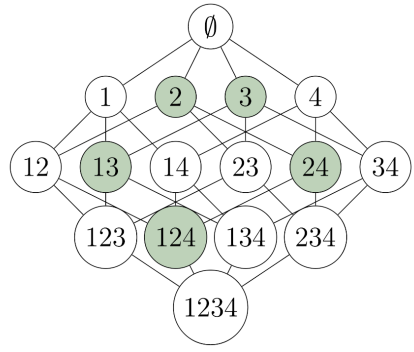
En raison de la nature combinatoire des interactions possibles, la puissance descriptive de ces modèles repose toutefois sur l'exploitation d'un grand nombre de paramètres, et leurs propriétés normatives, sur un grand nombre de contraintes. Par exemple, la capacité w paramétrant le modèle TD est définie par 2^n coefficients. Par conséquent, les méthodes *d'apprentissage de préférences* (appelées aussi *élicitation des préférences*), visant à adapter leurs paramètres au système de valeurs d'un décideur, s'appuient généralement sur des *restrictions préalables* de la flexibilité du modèle. Par exemple, une restriction largement utilisée consiste à autoriser uniquement les interactions

au sein de groupes de taille inferieure ou égale à un certain entier k, ou encore à spécifier en amont la collection $\mathcal{F}$ de groupes en interaction dans le modèle GAI [5, 9]. Cependant, bien que ces restrictions simplifient le problème d'apprentissage, elles amputent drastiquement le modèle d'une partie de sa richesse descriptive avant l'apprentissage.

Ainsi, exploitant les outils de l'apprentissage automatique et de l'optimisation pour l'apprentissage de modèles en haute dimension, nous proposons des algorithmes d'apprentissage de fonctions d'utilité avec interactions issues de la théorie de la décision, dont l'efficacité computationnelle ne repose pas sur des réductions préalables de la flexibilité des modèles (telles que, par exemple, la restriction des interactions possibles aux seules paires de points de vue, comme illustré sur le schéma (a) ci-dessous). En particulier, les algorithmes proposés réalisent un apprentissage *parcimonieux* des paramètres de ces fonctions d'utilité, permettant d'exploiter pleinement la richesse descriptive de ces modèles en laissant émerger, à partir des données de préférence, les interactions les plus importantes (comme illustré sur le schéma (b) ci-dessous, où les trois groupes $\{1,3\}$, $\{2,4\}$ et $\{1,2,4\}$, de cardinalité non limitée à 2, suffisent à décrire les préférences via les termes conjonctifs $\min(u_1(x_1), u_3(x_3))$, $\min(u_2(x_2), u_4(x_4))$ et $\min(u_1(x_1), u_2(x_2), u_4(x_4))$), produisant ainsi des modèles simples. Ce faisant, nous cherchons aussi à fournir à la communauté de l'apprentissage automatique des modèles de préférences interprétables et dotés de propriétés mathématiques souhaitables.



(a) Restriction des interactions aux groupe de taille k = 2.



(b) Émergence des interactions importantes sans condition sur la cardinalité.

Plus précisément, ces représentations sont apprises par la résolution de problèmes d'apprentissage régularisés par des fonctions favorisant les paramètres avec peu de coefficients non nuls. Ces fonctions de régularisation, dont

la norme ℓ_1 constitue un exemple canonique [18], se caractérisent par des points de non-différentiabilité là où les paramètres sont nuls, une propriété qui encourage les solutions parcimonieuses au cours du processus d'optimisation. D'un point de vue de l'optimisation, cette non-différentiabilité ajoute un niveau de difficulté supplémentaire à la tâche déjà complexe de l'optimisation dans des espaces de paramètres de haute dimension. Par conséquent, cette thèse exploite et adapte des paradigmes d'optimisation proposés pour résoudre de tels problèmes [2, 3, 4, 21], prenant notamment la forme d'algorithmes itératifs à faible coût par itération (mais à convergence lente) permettant d'obtenir des solutions approximatives au problème d'apprentissage dans des délais raisonnables, même pour des modèles impliquant des millions de paramètres.

Principales contributions

La donnée de départ des algorithmes d'apprentissage proposés est un ensemble d'exemples de préférences $x^j \succ x'^j$ et d'indifférences $x^j \sim x'^j$, ou d'évaluations numériques globales $x^j, y^j \in \mathbb{R}$, disponibles de manière différente selon le contexte. On considère notamment trois cadres d'apprentissage : (i) le cadre classique *en batch et passif* où l'on considère une base d'exemples précollectés, (ii) le cadre *en ligne* où les exemples sont accessibles sous la forme d'un flux incrémental, et finalement (iii) le cadre *actif* où les exemples sont sélectionnés et obtenus en interaction avec le décideur auquel on pose des questions.



Les algorithmes proposés, systématiquement accompagnés de garanties théoriques et validations empiriques dans la thèse, sont détaillés ci-dessous selon le modèle d'utilité considéré.

Modèles TD linéaires en w

Un algorithme dans le cadre en batch (i) est d'abord proposé pour l'intégrale de Choquet des utilités marginales, une instance centrale des modèles TD présentant l'avantange d'être linéaires en w. L'algorithme comprend l'estimation des utilités marginales et d'une représentation parcimonieuse de la capacité w via sa *transformée de Möbius* [10, 13]. Les deux étapes reposent sur des programmes linéaires, résolubles à l'exact à l'aide de solveurs numériques. Ils ne sont cependant pas résolubles à l'exact pour plus d'une douzaine de points de vue du fait d'un nombre de variables et de contraintes exponentiel en n. Ainsi, dans un cadre unifié pour l'apprentissage des modèles TD linéaires en w, incluant notamment l'intégrale de Choquet et le modèle multilinéaire, nous proposons une méthode d'optimisation computationnellement efficace, fondée sur le principe des *moindres carrés itérés repondérés* et la *dualité Lagrangienne* [11]. Enfin, nous proposons dans le cadre en ligne (ii), une méthode de mise à jour computationnellement efficace de ces modèles à chaque arrivée de nouvel example, en exploitant le principe des méthodes *follow-the-regularized-leader*, que l'on combine ensuite avec le principe de la *méthode des multiplicateurs à directions alternées* pour inclure des contraintes structurelles sur la capacité w telles que la monotonie par rapport à l'inclusion [12].

Modèles TD sans condition de linéarité en w

Un algorithme général d'apprentissage, adapté à toute fonction d'agrégation paramétrique F_w (à utilités marginales connues), est ensuite proposé dans le cadre actif (iii). Basé sur le principe des méthodes d'*apprentissage actif par désagrément*, il permet de réduire progressivement l'espace des paramètres possibles en extrayant du décideur les exemples les plus informatifs, jusqu'à ce que l'incertitude concernant la solution préférée dans un ensemble de solutions de départ soit suffisamment réduite pour pouvoir lui formuler une recommandation [14]. L'algorithme permet, de manière novatrice, de tenir compte du bruit dans les réponses du décideur, tout en évitant la lourdeur computationnelle des méthodes Bayesiennes.

Modèles GAI

Un algorithme d'apprentissage de fonctions GAI est finalement proposé dans le cadre batch (i). Il est capable simultanément, d'identifier les facteurs correspondant aux attributs en interaction et d'apprendre les fonctions d'utilité définies sur ces facteurs, sans restreindre la taille des groupes d'interactions possibles. L'approche proposée repose sur la détermination d'une représentation parcimonieuse de la décomposition *ANOVA* de la fonction d'utilité, en utilisant l'*apprentissage par noyaux multiples* [15].

Perspectives

Ces travaux ouvrent la voie à l'apprentissage de modèles de préférences de formes variées, allant au delà de la fonction d'utilité qui présuppose certaines propriétés structurelles des préférences telles que la *complétude* (toutes les solutions sont comparables) et la *transitivité* (si x est préféré à y et y à z, alors x est préférée à z).

Les algorithmes proposés pourraient en effet être étendus à l'apprentissage de modèles potentiellement non transitifs, par exemple de la forme $P_w(x, x') = F_w(u_1(x_1) - u_1(x'_1), \ldots, u_n(x_n) - u_n(x'_n))$, où chaque alternative x ou x' se voit associer un score décrivant l'intensité de la préférence de x sur x', obtenu en agrégeant les écarts de performance entre x et x' selon les différents points de vue (i.e. $u_i(x_i) - u_i(x'_i),\ i = 1, \ldots, n$). Ceci permettrait d'accéder à une palette encore plus riche de représentations de préférences, tout en maintenant les objectifs d'efficacité computationnelle et d'interprétabilité des modèles atteints dans cette thèse.

Un autre prolongement naturel de ces travaux consiste en leur intégration au sein de systèmes d'intelligence artificielle opérationnels, tels que les *systèmes de recommandation* ou les *agents conversationnels*, afin d'assurer la traçabilité et l'interprétabilité des décisions prises ou recommandées par ces systèmes, et de les doter de certaines propriétés normatives jouant le rôle de garde-fous, dans la perspective d'une intelligence artificielle *transparente* et *rationnelle*.

Références

[1] F. Bacchus et A. Grove. 1995. Graphical Models for Preference and Utility. In *UAI'95*.

[2] F. Bach, R. Jenatton, J. Mairal, G. Obozinski, et al. 2012. Optimization with sparsity-inducing penalties. *Foundations and Trends in Machine Learning* 4, 1: 1-106.

[3] A. Beck et M. Teboulle. 2009. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM journal on imaging sciences* 2, 1: 183-202.

[4] L. Bottou, F. E. Curtis, et J. Nocedal. 2018. Optimization Methods for Large-Scale Machine Learning. *SIAM Review* 60, 2: 223-311. https://doi.org/10.1137/16M1080173.

[5] R. Bresson, J. Cohen, E. Hüllermeier, C. Labreuche, et M. Sebag. 2021. Neural representation and learning of hierarchical 2-additive Choquet integrals. In *Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence*, 1984-1991.

[6] H. Doukas et A. Nikas. 2020. Decision support models in climate policy. *European Journal of Operational Research* 280, 1: 1-24.

[7] P. C. Fishburn. 1970. *Utility Theory for Decision Making*. Wiley.

[8] M. Grabisch. 1996. The application of fuzzy integrals in multicriteria decision making. *European Journal of Operational Research* 89, 3: 445-456.

[9] M. Grabisch, C. Labreuche, et M. Ridaoui. 2022. Well-formed decompositions of Generalized Additive Independence models. *Annals of Operations Research* 312, 2: 827-852.

[10] M. Herin, P. Perny, et N. Sokolovska. 2022. Learning Sparse Representations of Preferences within Choquet Expected Utility Theory. In *The 38th Conference on Uncertainty in Artificial Intelligence*.

[11] M. Herin, P. Perny, et N. Sokolovska. 2023. Learning Preference Models with Sparse Interactions of Criteria. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence, IJCAI-23*, 3786-3794.

[12] M. Herin, P. Perny, et N. Sokolovska. 2024. Online Learning of Capacity-Based Preference Models. In *International Joint Conference on Artificial Intelligence (IJCAI) 2024*, 7118-7126.

[13] M. Herin, P. Perny, et N. Sokolovska. 2024. Learning preference representations based on Choquet integrals for multicriteria decision making. *Annals of Mathematics and Artificial Intelligence*: 1-34.

[14] M. Herin, P. Perny, et N. Sokolovska. 2024. Noise-Tolerant Active Preference Learning for Multicriteria Choice Problems. In *International Conference on Algorithmic Decision Theory*, 191-206.

[15] M. Herin, P. Perny, et N. Sokolovska. 2024. Learning GAI-decomposable Utility Models for Multiattribute Decision Making. In *The 38th Annual AAAI Conference on Artificial Intelligence (AAAI 2024)*. Consulté à l'adresse `https://hal.science/hal-04424705`

[16] G. Owen. 1972. Multilinear extensions of games. *Management Science* 18, 5-part-2: 64-79.

[17] E. Pohl et J. Geldermann. 2024. Selection of multi-criteria energy efficiency and emission abatement portfolios in container terminals. *European Journal of Operational Research* 316, 1: 386-395.

[18] R. Tibshirani. 1996. Regression Shrinkage and Selection via the lasso. *Journal of the Royal Statistical Society. Series B (methodological)* 58, 1: 267-88.

[19] A. Tversky et D. Kahneman. 1981. The framing of decisions and the psychology of choice. *science* 211, 4481: 453-458.

[20] J. von Neumann et O. Morgenstern. 1944. Theory of games and economic behavior.

[21] L. Xiao. 2010. Dual Averaging Methods for Regularized Stochastic Learning and Online Optimization. *Journal of Machine Learning Research* 11: 2543-2596.