

# Aide à la décision multicritère et apprentissage automatique pour la classification

Antoine Rolland\*

\* Laboratoire ERIC - Université Lumière Lyon II  
av Pierre Mendès-France, 69676 BRON Cedex  
antoine.rolland@univ-lyon2.fr

**Résumé.** Les méthodes d'aide multicritère à la décision pour la problématique du tri d'une part, et les méthodes d'apprentissage automatique d'autre part poursuivent le même objectif : permettre d'affecter des objets à des catégories prédéfinies, ordonnées ou non. Cependant, les différences existantes entre les deux domaines d'application font que jusqu'à aujourd'hui ces deux domaines ont eu peu d'interactions. Il nous semble cependant qu'il y a la matière à exploration. Nous proposons donc de présenter quelques méthodes et paradigmes pour la problématique du tri en analyse multicritère avant de proposer quelques pistes pour des applications croisées en apprentissage automatique et en aide à la décision.

## 1 Introduction

Prendre une décision, c'est trancher entre plusieurs possibilités et choisir celle qui sera effectivement mise en œuvre. Les différentes communautés qui se sont intéressées à la prise de décision (économie, sociologie, psychologie, mathématiques, statistique, informatique...) se sont toutes penchées d'une manière ou d'une autre sur cette question, pour essayer de comprendre comment l'être humain prend une décision, pour l'accompagner dans cette prise de décision ou pour pouvoir reproduire cette décision de manière automatique. Prendre une décision est un acte simple quand toutes les alternatives sont connues, qu'elles sont peu nombreuses, qu'elles peuvent être évaluées de manière unique, et qu'il n'y a qu'une seule personne qui décide. Cependant, si les alternatives ou leurs conséquences sont imparfaitement connues, si leur nombre est trop important pour qu'une approche systématique puisse être envisagée, si elles doivent être évaluées selon plusieurs critères partiellement conflictuels, ou si plusieurs personnes sont amenées à se prononcer, alors la prise de décision devient un acte complexe. Chacune de ces difficultés a donné naissance à un champ particulier de recherche : décision dans l'incertain, optimisation combinatoire, décision multicritère, décision multi-agents... A la suite de Roy (1996), on distingue souvent trois types de problématique en aide à la décision :

- la problématique du choix, qui consiste à vouloir *choisir* la ou les solutions considérées comme optimales pour le problème considéré ;
- la problématique du classement (ranking), qui consiste à vouloir classer du premier au dernier toutes les solutions connues du problème considéré ;
- la problématique du tri (sorting), qui consiste à affecter les solutions à des catégories (ordonnées ou non).

Nous nous intéresserons ici au cas où les solutions étudiées sont évaluées suivant plusieurs points de vue, ou critères, qui peuvent potentiellement être conflictuels. Nous nous focaliserons en particulier sur la problématique du tri dans des catégories ordonnées.

### 1.1 Paradigme de l'aide à la décision multicritère

En aide à la décision multicritère (MCDA ou MultiCriteria Decision Aiding), on suppose que les différentes alternatives se présentant au décideur peuvent être décrites sur un certain nombre de propriétés ou attributs. Si l'on est capable d'établir une échelle de préférence sur un attribut, on parle alors de critère. Les valeurs des alternatives sur ces critères représentent la prise en compte de points de vue diversifiés, en général non réductibles à un seul critère. Ces critères sont souvent conflictuels. La difficulté est d'arriver à obtenir une comparaison relative des alternatives, afin de guider le choix du décideur vers la solution qui lui paraîtra optimale. En effet, la notion d'optimisation "dans l'absolu" est vide de sens en décision multicritère, car il n'existe généralement pas d'alternative optimisant tous les critères simultanément. Il est donc nécessaire de prendre en compte de l'information supplémentaire, en particulier l'importance relative de chaque critère et les compensations possibles entre les différents critères.

Le but des méthodes en MCDA est donc moins de proposer automatiquement une unique solution considérée comme optimale, que d'accompagner le décideur dans sa démarche de décision. La phase de définition du problème, des alternatives et des critères est généralement considérée comme plus importante que la phase purement calculatoire mettant en œuvre une méthode particulière d'agrégation de préférence (voir Schärli (1985), Vincke (1992), Roy (1996)). En particulier, l'interaction avec le décideur tout au long du processus fait que le résultat d'une méthode d'aide à la décision ne semble pas arriver de nulle part, mais permet bien au décideur de prendre sa décision de manière éclairée.

### 1.2 Aide à la décision et apprentissage automatique

Il existe de nombreuses méthodes d'aide multicritère à la décision. Chaque méthode met en jeu un certain nombre de paramètres. Quelle que soit la méthode d'aide à la décision retenue, la question de fixer les valeurs des paramètres de la méthode est une question cruciale en aide à la décision. Deux approches existent à ce propos :

- demander au décideur de fixer les valeurs des paramètres directement : cette voie est délicate, du fait que le décideur ne maîtrise pas l'effet que peut avoir tel ou tel paramètre sur les comparaisons d'alternatives, et que les concepts utilisés dans chacune des méthodes peuvent parfois être difficile à appréhender pour un non-expert (par exemple la différence entre les poids des critères dans une méthode de moyenne pondérée et les degrés d'importance des critères dans une méthode ELECTRE n'est pas évidente).
- retrouver les valeurs des paramètres de la méthode à partir d'un jeu de données existant. Il s'agit d'inférer les paramètres de la méthode en prenant comme corpus d'apprentissage un jeu de données sur lequel on recueille les préférences du décideur. C'est cette méthode, proche dans son esprit de l'apprentissage automatique, que nous allons étudier par la suite, sous le nom d'élicitation des paramètres.

Cependant, il existe de grandes différences d'approches entre l'apprentissage automatique et l'élicitation des préférences en MCDA. Waegeman et al. (2009) détaillent huit principaux points de divergence que nous reprenons dans la table 1. De manière générale, on peut résumer

TAB. 1 – Principales différences entre aide multicritère à la décision et apprentissage automatique, selon Waegeman et al. (2009)

Thème	MCDA	Apprentissage automatique
Information disponible	orientée données	orientée résultat
Interaction avec le décideur	oui	non
Décideur	peut changer d'avis	préférences fixes
Cohérence	données cohérentes	bruit possible
Taille de l'échantillon	petit	grand
Modèle	souvent additif	très divers
Régularisation	non	oui
Adaptabilité	mauvaise	bonne
Données	vectérielles	de toute sorte

cela en disant que l'apprentissage automatique travaille avec des données très diverses, sur lesquelles on ne peut agir, disponibles dans des corpus de grande taille mais pouvant contenir des erreurs. Au contraire, les méthodes d'élicitation en MCDA utilisent généralement des données vectorielles provenant de corpus de petite taille, sans incohérence, et modifiables par dialogue avec le décideur. Malgré ces différences, nous pensons que les méthodes en MCDA peuvent être intéressantes également dans le cadre de l'apprentissage automatique. Nous allons donc succinctement présenter ces méthodes et les techniques d'élicitation usuelles (section 2) puis nous indiquerons des pistes de recherche pour rapprocher les deux domaines (section 4)

## 2 Méthodes multicritères pour le tri

Les méthodes de tri en MCDA s'intéressent à des objets appelés alternatives, décrites par plusieurs attributs, aussi appelés critères s'ils sont dotés d'une relation binaire de préférence. Le but de ces méthodes est d'affecter chaque alternative à l'un des catégories prédéfinies, qu'elles soient ordonnées ou non (Roy (1996)). Formellement, nous allons considérer ici un problème de tri en aide à la décision multicritère où les alternatives de l'ensemble  $A = \{a_1, a_2, \dots, a_j, \dots, a_m\}$  sont évaluées sur  $n$  critères  $g_1, g_2, \dots, g_i, \dots, g_n$ , avec  $i \in N = \{1, \dots, i, \dots, n\}$ . Le critère  $g_i$  prend ses valeurs sur un ensemble  $X_i$ , i.e.,  $g_i : A \mapsto X_i$ , doté d'une relation de préférence  $\succsim_i$ .  $X$  est alors le produit cartésien des différentes échelles de critère ( $X = \prod_{i \in N} X_i$ ). Toute alternative de  $A$  peut donc être décrite par un vecteur de  $X$ . Les catégories ordonnées prédéfinies forment une partition de  $A$  et sont notées  $C_1, C_2, \dots, C_p$  où  $C_i$  est préférée à  $C_j$  si  $i > j$ . On définit  $C_t^{\geq}$ ,  $t = 1, \dots, p$ , par  $C_t^{\geq} = \cup_{s \geq t} C_s$ .

Il existe aujourd'hui un large panel de méthodes d'aide à la décision multicritère. On se référera à Bouyssou et al. (2006) pour une vue complète des méthodes les plus connues, et à Zopounidis et Doumpos (2002) pour les méthodes de tri. On distingue généralement trois grands types de méthodes pour l'aide à la décision multicritère, en reprenant la terminologie de Roy (1996) pour les deux premières :

- les méthodes à score unique de synthèse : il s'agit, pour chaque alternative, d'agréger toutes les valeurs prises sur chacun des critères pour obtenir un "score" unique pour cha-

cune des alternatives, un critère unique de synthèse . Il suffit ensuite pour comparer deux alternatives de comparer leurs scores respectifs. Ces méthodes ont été historiquement développées sous le nom de MultiAttribute Utility Theory (MAUT) depuis les années 70 principalement aux Etats-Unis (voir Keeney et Raiffa (1976)). L'introduction d'utilité non-additive a permis de généraliser ces méthodes. On consultera Marichal (2009) pour une vue d'ensemble de ces méthodes formelles. Dans le cadre particulier des méthodes de tri, on peut mentionner la méthode UTADIS proposée par Jacquet-Lagrèze (1995).

- les méthodes à relation de synthèse : il s'agit ici au contraire de comparer, critère par critère, les deux alternatives afin d'obtenir autant de relations de préférence partielles entre les deux alternatives qu'il existe de critères, puis d'essayer d'agréger ces préférences partielles en une préférence globale. Ces méthodes, basées sur une représentation des préférences par des graphes, a été développée depuis 1975, en particulier en Europe et dans la communauté francophone.
- plus récemment, les méthodes utilisant les règles de décision et le principe de dominance ont été proposées par Greco et al. (2001b).

Afin d'illustrer ces trois approches, et devant la multiplicité des méthodes multicritères existantes, nous choisissons de présenter ci-dessous trois méthodes "emblématiques" pour les problèmes de tri en MCDA. La première méthode, utilisant une fonction d'utilité non-additive (intégrale de Choquet). La deuxième, utilise une comparaison par paire (ELECTRE TRI), et la troisième les règles de décision.

## 2.1 Fonction d'utilité non-additive utilisant l'intégrale de Choquet

L'intégrale de Choquet est une méthode de scoring donnant une valeur pour chaque alternative à partir d'une mesure non additive sur les ensembles de critères. Une mesure non-additive, aussi appelée capacité (voir Grabisch (1996), Grabisch et Roubens (2000), Marichal (2009)), est une fonction  $v : 2^N \rightarrow \mathbb{R}^+$  telle que  $v(\emptyset) = 0$ , et pour deux sous-ensembles  $A$  et  $B$  de  $N$ ,  $A \subseteq B \subseteq N$  implique  $v(A) \leq v(B)$  (principe de monotonie par inclusion). On normalise généralement la capacité par  $v(\emptyset) = 0$  et  $v(N) = 1$ . Une mesure non-additive permet par exemple que  $v(A \cup B) \neq v(A) + v(B)$ . L'intégrale de Choquet  $u(a)$  est alors définie par :

$$u(a) = \sum_{i=1}^n g_{\tau(i)}(a)(v(\{\tau(i), \dots, \tau(n)\}) - v(\{\tau(i+1), \dots, \tau(n)\}))$$

où  $\tau$  est une permutation sur  $N$  telle que  $g_{\tau(1)}(a) \leq g_{\tau(2)}(a) \leq \dots \leq g_{\tau(n)}(a)$ . L'intégrale de Choquet est un opérateur d'agrégation très général, pouvant modéliser les opérateurs additifs ou non-additif tels que le min, le max, la somme pondérée, OWA, etc (voir Grabisch et al. (2009) et Grabisch et Labreuche (2008) pour les utilisations de l'intégrale de Choquet). L'utilisation d'une intégrale de Choquet nécessite que les ensembles  $X_i$  soient commensurables, par exemple  $[0, 1] \forall i \in N$ . Formellement, l'intégrale de Choquet est une méthode de classement (ranking method), mais elle peut être utilisée comme méthode de tri (sorting method) en définissant un niveau de référence  $\lambda_t$  par catégorie  $C_t$  tel que :

$$a \in C_t^{\geq} \iff u(a) \geq \lambda_t$$

Comme toutes les méthodes à base d'utilités, la méthode de tri utilisant l'intégrale de Choquet repose sur plusieurs hypothèses fortes. En particulier, on suppose :

- la commensurabilité des différentes échelles de valeurs sur les critères entre elles
- la commensurabilité des échelles de valeurs sur les critères et la mesure  $v$  choisie
- la connaissance complète de la mesure  $v$ , et en particulier de la compensation intercritères.

Ces hypothèses permettent un calcul facile de la relation de préférence entre les différentes alternatives, et en particulier de l'affectation des alternatives aux différentes catégories. Mais cette modélisation est très couteuse en information sur les paramètres de la méthode. De plus, la facilité des calculs est obtenue au prix d'un éloignement certain de la manière dont raisonne un décideur.

## 2.2 ELECTRE-Tri

Une deuxième approche utilise une méthode à base de surclassement appelée ELECTRE-Tri (Roy (1991), Figueira et al. (2005)). Cette méthode utilise des profils prédéterminés, éléments de  $X$  correspondant à des "alternatives fictives", pour séparer les différentes catégories. Une alternative est alors affectée à la catégorie  $C_i$  (ou supérieure) si elle est préférée au profil marquant la limite inférieure de la catégorie. La relation de préférence utilisée dans ELECTRE-Tri est une relation de surclassement utilisant la concordance et la non-discordance comme dans les méthodes ELECTRE en général (Roy (1996)). De manière formelle, nous avons  $p - 1$  profils  $q^2, \dots, q^p$  prenant leurs valeurs sur  $X$ ,  $n$  indices d'importance des critères  $\omega_1, \dots, \omega_n$  et  $\lambda$  un niveau seuil tel que :  $a$  est affecté à la catégorie  $C_i$  ou supérieure si

$$\sum_{i|g_i(a) \geq p_i^t} \omega_i \geq \lambda$$

Cette approche est moins gourmande en informations initiales, supporte des relations purement ordinales sur les valeurs des critères, voire des données imprécises ou manquantes. A contrario, les résultats obtenus sont généralement moins discriminants que les méthodes à base d'utilités. En particulier, des phénomènes d'incomparabilité entre alternatives et profils peuvent survenir, aboutissant à une incertitude possible sur les affectations à des catégories (voir Roy (1968, 1996)).

## 2.3 Règles de décision

Une troisième approche, proposée par Greco et al. (2001b) et Greco et al. (2002), utilise les règles de décision à travers les ensembles approximatifs et le principe de dominance (Dominance-based Rough Set Approach). On dit qu'une alternative  $a$  domine une alternative  $b$  si  $\forall i \in N, g_i(a) \succsim_i g_i(b)$ . Les alternatives sont affectées aux différentes catégories en fonction de niveaux de référence sur chaque critère. Les fondements axiomatiques de ces méthodes ont été étudiés, en particulier dans le cadre de la problématique du tri, par Greco et al. (2001a) pour les relations de surclassement. Formellement, un modèle à base de règles de décision est capable d'affecter l'alternative  $a$  à une catégorie grâce à des règles du type "si  $g_1(a) \succsim_1 \alpha_1$  et  $g_2(a) \succsim_2 \alpha_2$  et  $\dots$  alors  $a$  est affecté à la catégorie  $C_i$ ".

Le grand intérêt de ce type de méthode réside dans l'interprétation immédiate des résultats et la représentation sous forme de règles de la logique du décideur. Cela se fait au détriment d'un modèle synthétique : en effet, le nombre de règles de décision peut rapidement être trop important pour être appréhendé rapidement.

### 3 Elicitation des paramètres des méthodes de tri

Récemment, plusieurs auteurs se sont particulièrement intéressés aux processus d'élicitation pour les méthodes multicritères d'aide à la décision, tant en étudiant les fondements théoriques de ces processus qu'en proposant des outils concrets pour l'élicitation. L'idée fondatrice est de partir d'exemples d'affectations données par le décideur pour calculer les paramètres de la méthode choisie redonnant ces affectations. On a donc en entrée de la procédure d'élicitation l'ensemble des alternatives et de leurs valeurs sur chacun des critères, généralement présenté sous forme d'un tableau appelé table de performance, ainsi que l'affectation de chacune des alternatives à une catégorie.

Dans le cadre de l'intégrale de Choquet, des procédures d'estimation des coefficients d'une capacité ont été proposées par Grabisch et al. (2008). Le point de départ de ces procédures est le suivant : connaissant la table de performance des alternatives sur tous les critères et le score global de chaque alternative, le but est de trouver les paramètres de la capacité qui fait en sorte que les résultats du calcul de l'intégrale de Choquet soient les plus proches possibles de ces valeurs globales. Les méthodes retenues utilisent la programmation linéaire avec plusieurs fonctions objectifs possibles, comme les moindres carrés ou la variance minimale. Le package `kappalab` permettant l'élicitation des capacités est disponible sous 'R' et dans le projet Decision Deck (Meyer et Bigaret (2011)) proposé par Meyer (2011).

L'avantage de ces méthodes est de toujours aboutir à une solution ; l'inconvénient est que la solution proposée peut être loin de retrouver la classification initiale si certaines incohérences existent chez le décideur. En outre, la complexité des calculs fait que le programme ne fonctionne que dans les cas où ni les critères ni les alternatives ne sont trop nombreuses.

Les procédures d'élicitation des paramètres de la méthode ELECTRE-Tri ont été proposées par Mousseau et Slowinski (1998) (voir aussi Ngo The et Mousseau (2002), Mousseau et al. (2001) et Bouyssou et Marchant (2007a,b)). Ces procédures utilisent un programme linéaire pour minimiser le nombre de différence entre les affectations aux catégories proposées par le décideur d'une part et par la méthode ELECTRE-TRI d'autre part. Le nombre de profils est déterminé par le nombre de catégories. Une implémentation de cette procédure dans le projet Decision Deck a été proposée par Sobrie (2011).

La encore, la taille de l'ensemble d'apprentissage est un facteur limitant l'efficacité de l'approche. Cependant, la possibilité de pouvoir fixer *a priori* certains des paramètres tels que la valeur de seuil permet potentiellement d'accélérer les calculs.

La méthode d'apprentissage des règles de décision utilise une approche fondée sur les ensembles approximatifs et la relation de dominance à partir d'un ensemble d'apprentissage, comme proposé par Greco et al. (2000). On dit que l'alternative  $a$  domine l'alternative  $b$  si  $g_i(a) \succsim_i g_i(b)$  pour tout critère  $g_i, i \in N$ . Pour un ensemble de critères  $\{g_i, i \in N\}$  donné, l'inclusion d'une alternative  $a \in A$  dans l'union des classes  $C_t^{\succsim}, t = 2, \dots, p$  aboutit à une incohérence au sens du principe de dominance si une des deux conditions suivantes tient :

- $a$  appartient à la catégorie  $C_t^{\succsim}$  ou une catégorie supérieure, mais est dominée par une alternative  $b$  appartenant à une catégorie moins élevée que  $C_t$
- $a$  appartient à une catégorie inférieure à  $C_t^{\succsim}$  mais domine une alternative  $b$  appartenant à la catégorie  $C_t$  ou une catégorie supérieure

Si l'inclusion de  $a \in A$  dans  $C_t^{\succsim}, t = 2, \dots, p$ , amène une incohérence au sens du principe de dominance, alors  $a$  appartient à  $C_t^{\succsim}$  avec ambiguïté. Et donc  $a$  appartient à  $C_t^{\succsim}$  sans ambiguïté

si  $a \in C_t^{\geq}$  et s'il n'y a aucune incohérence au sens du principe de dominance. Cela signifie que toutes les alternatives dominant  $a$  appartiennent à  $C_t^{\geq}$ . Les affectations des alternatives de l'ensemble d'apprentissage peuvent donc être séparées entre affectations certaines et affectations possibles. Ces affectations entraînent ensuite des règles de décision certaines et des règles de décision possibles. L'affectation d'une nouvelle alternative se fait alors à l'aide des règles de décision certaines. le logiciel jMAF proposé par Stefanowski (1998) permet la génération des règles de décision à partir d'un corpus d'apprentissage.

Le grand avantage de cette méthode est de pouvoir traiter des incohérences dans les affectations proposées par le décideur. l'inconvénient est que cela peut aussi parfois aboutir à un grand nombre de règles de décision, chacune traitant un cas spécifique.

## 4 Perspectives de recherche

Les différences évoquées entre apprentissage automatique et MCDA nous conduisent à proposer quatre pistes de recherche où les deux champs pourraient interagir. L'objectif poursuivi est l'intégration de méthodes provenant de l'aide à la décision dans le domaine de l'apprentissage automatique. En particulier, un travail d'investigation doit être mené pour l'identification de champs d'application pour lesquelles l'utilisation des méthodes d'aide à la décision pourraient apporter une réelle plus-value.

1. une première piste de recherche concerne la **complexité des calculs** pour l'élicitation des paramètres. Les algorithmes utilisés reposent souvent sur des algorithmes de programmation linéaire, dont le temps de calcul s'accroît rapidement avec le nombre de variables en entrée (critère et/ou instances). A titre d'exemple, le calcul des capacités par Kappalab est limité à moins de 10 critères, ce qui est amplement suffisant en aide à la décision (il est courant de considérer qu'un problème à plus de 6 ou 7 critère est un problème mal défini) mais pas forcément en apprentissage automatique. Une prise en compte de la structure particulière des problèmes d'apprentissage automatique permettrait peut-être de réduire le temps de calcul, en particulier pour les méthodes du type "utilité" ou "ELECTRE Tri".
2. une deuxième piste a trait au **traitement des erreurs et incohérences** : la différence de point de vue entre l'apprentissage automatique, où les erreurs sont acceptées comme telles, et l'aide à la décision, où l'on suppose que les préférences indiquées sont cohérentes, fait que les méthodes de MCDA visent à intégrer toutes les données disponibles pour éliciter les paramètres, ce qui est inefficace dans le cas de grands corpus d'apprentissage. Il nous semble qu'une piste de recherche pourrait s'intéresser à la robustesse des algorithmes d'élicitation vis-à-vis de la présence ou de l'absence de telle ou telle alternative dans le corpus d'apprentissage, afin de pouvoir éliminer les "valeurs aberrantes".
3. une troisième piste consiste à pouvoir **interagir avec le décideur** dans le cadre de l'apprentissage automatique. cela semble paradoxal avec le terme "automatique" mais une interaction avec le décideur pourrait, en complément du point précédent, permettre aux méthodes d'élicitation de ne pas prendre en compte des données pouvant être considérées comme erronées. Une autre idée consiste à permettre au décideur de réduire l'ensemble d'apprentissage à quelques cas emblématiques pour que la taille du corpus soit compatible avec les capacités de calcul des méthodes d'élicitation. La question est alors

celle de l'analyse du corpus de données en entrée pour ôter l'information redondante et ne garder que les alternatives vraiment intéressantes.

4. enfin, une quatrième piste concerne l'intégration d'outil d'aide à la décision pour **l'agrégation de classifieurs**, dans le cas où l'apprentissage automatique se fait à travers plusieurs algorithmes potentiellement conflictuels (cf Kuncheva (2004), Stefanowski et Nowaczyk (2007)). L'idée est d'agir non pas directement sur le corpus des alternatives, mais sur les résultats obtenus à l'aide de différents classifieurs, soit faisant partie d'une même famille de classifieurs avec des paramètres différents, soit provenant de plusieurs approches différentes de classifieurs.

## Références

- Bouyssou, D., D. Dubois, M. Pirlot, et H. Prade (2006). *Concepts et méthodes pour l'aide à la décision, volume 3, analyse multicritère*. Hermes.
- Bouyssou, D. et T. Marchant (2007a). An axiomatic approach to noncompensatory sorting methods in mcdm, I : The case of two categories. *European Journal of Operational Research* 178, 217–245.
- Bouyssou, D. et T. Marchant (2007b). An axiomatic approach to noncompensatory sorting methods in mcdm, II : More than two categories. *European Journal of Operational Research* 178, 246–276.
- Figueira, J., V. Mousseau, et B. Roy (2005). ELECTRE methods. In J. Figueira, S. Greco, et M. Ehrgott (Eds.), *Multiple Criteria Decision Analysis : State of the Art Surveys*, pp. 133–162. Boston, Dordrecht, London : Springer Verlag.
- Grabisch, M. (1996). The application of fuzzy integrals in multicriteria decision making. *European Journal of Operational Research* 89, 445–456.
- Grabisch, M., I. Kojadinovic, et P. Meyer (2008). A review of methods for capacity identification in choquet integral based multi-attribute utility theory : Applications of the kappalab r package. *European Journal of Operational Research* 186(2), 766 – 785.
- Grabisch, M. et C. Labreuche (2008). A decade of application of the choquet and sugeno integrals in multi-criteria decision aid. *4OR : A Quarterly Journal of Operations Research* 6, 1–44.
- Grabisch, M., J.-L. Marichal, R. Mesiar, et E. Pap (2009). *Aggregation functions*, Volume 127 of *Encyclopedia of Mathematics and its Applications*. Cambridge, UK : Cambridge University Press.
- Grabisch, M. et M. Roubens (2000). Application of the Choquet integral in multicriteria decision making. In M. Grabisch, T. Murofushi, et M. Sugeno (Eds.), *Fuzzy Measures and Integrals - Theory and Applications*, pp. 348–374. Physica Verlag.
- Greco, S., B. Matarazzo, et R. Slowinski (2000). Multicriteria classification by dominance-based rough set approach, methodological basis of the 4emka system. [Online], Available : <http://www-idss.cs.put.poznan.pl/4emka/>.
- Greco, S., B. Matarazzo, et R. Slowinski (2001a). Conjoint measurement and rough set approach for multicriteria sorting problems in presence of ordinal criteria. In A. Colorni,



- M. Paruccini, et B. Roy (Eds.), *A-MCD-A, Aide Multicritère à la Décision/Multiple Criteria Decision Aid*, pp. 117–144. Ispra : European Commission, Joint Research Centre, EUR 19808 EN.
- Greco, S., B. Matarazzo, et R. Slowinski (2001b). Rough sets theory for multicriteria decision analysis. *European Journal of Operational Research* 129, 1–47.
- Greco, S., B. Matarazzo, et R. Slowinski (2002). Rough sets methodology for sorting problems in presence of multiple attributes and criteria. *European Journal of Operational Research* 138, 247–259.
- Jacquet-Lagrèze, E. (1995). An application of the UTA discriminant model for the evaluation of sc R&D projects. In P. Pardalos, Y. Siskos, et C. Zopounidis (Eds.), *Advances in Multi-criteria Analysis, Nonconvex Optimization and its Applications*, pp. 203–211. Dordrecht : Kluwer Academic.
- Keeney, R. et H. Raiffa (1976). *Decisions with multiple objectives : Preferences and value tradeoffs*. J. Wiley, New York.
- Kuncheva, L. (2004). *Combining Pattern Classifiers. Methods and Algorithms*. Wiley.
- Marichal, J.-L. (2009). *Aggregation functions for decision making, Decision-Making Process - Concepts and Methods*. ISTE/John Wiley.
- Meyer, P. (2011). <http://www.decision-deck.org/ws/wsd-choquetintegral-kappalab.html>.
- Meyer, P. et S. Bigaret (2011). diviz : a software for modeling, processing and sharing algorithmic workflows in mcda. *Intelligent Decision Technologies : an International Journal*, to appear.
- Mousseau, V., J. Figueira, et J. Naux (2001). Using assignment examples to infer weights for ELECTRE TRI method : Some experimental results. *European Journal of Operational Research* 130(2), 263–275.
- Mousseau, V. et R. Slowinski (1998). Inferring an electre tri model from assignment examples. *Journal of Global Optimization* 12, 157–174.
- Ngo The, A. et V. Mousseau (2002). Using assignment examples to infer category limits for the ELECTRE TRI method. *JMCDA* 11(1), 29–43.
- Roy, B. (1968). Classement et choix en présence de point de vue multiples (la méthode electre). *Les cahiers du CERO* (8), 57–75.
- Roy, B. (1991). The outranking approach and the foundations of ELECTRE methods. *Theory and Decision* 31, 49–73.
- Roy, B. (1996). *Multicriteria Methodology for Decision Aiding*. Dordrecht : Kluwer Academic.
- Schärli, A. (1985). *Décider sur plusieurs critères*. Presses Polytechniques Universitaires Romandes.
- Sobrie, O. (2011). <http://www.decision-deck.org/ws/wsd-electretribminference-pyxmcda.html>.
- Stefanowski, J. (1998). On rough set based approaches to induction of decision rules. In S. A. Polkowski L. (Ed.), *Rough Sets in Data Mining and Knowledge Discovery*, Volume 1, pp. 500–529. Physica-Verlag.
- Stefanowski, J. et S. Nowaczyk (2007). An experimental study of using rule induction algo-

MCDA et apprentissage automatique

rithm in combiner multiple classifier. *International Journal of Computational Intelligence Research* 3, 335–342.

Vincke, P. (1992). *Multicriteria Decision-Aid*. J. Wiley, New York.

Waegeman, W., B. D. Baets, et L. Boullart (2009). Kernel-based learning methods for preference aggregation. *4OR* (7), 169–189.

Zopounidis, C. et M. Doumpos (2002). Multicriteria classification and sorting methods : A literature review. *European Journal of Operational Research* (138), 229–246.

## Summary

Multicriteria decision aiding methods for sorting problems and machine learning algorithms are sharing a similar goal : they aim at assigning objects to one of the predefined ordered or not ordered categories. However, the differences between these two fields of application avoid seminal interactions. It seems to us that it could be interesting to study how they can interact. In this paper, we present some methods and paradigms for sorting multicriteria decision aiding methods, and then we suggest some tracks for further research.