

Stratégies d'auto détection de modèles erronés en élicitation de données personnalisées dans un environnement incertain

Laboratoire : LIAS (<https://www.lias-lab.fr>)
Établissement : ISAE-ENSMA (<https://www.ensma.fr>)
Encadrants : Loïc ADAM, Allel HADJALI
Durée : Jusqu'à 6 mois

1 Contexte

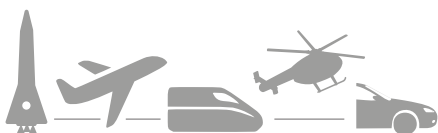
Ce projet de recherche a pour objectif de mêler *élicitation de profils utilisatrices*, sous forme de *préférences* (en décision multicritère), *gestion de l'incertitude* et de Machine Learning (ML). L'élicitation de préférences cherche à aider une utilisatrice à prendre une décision parmi un large choix d'alternatives (choisir un plan d'investissement, par exemple), tout en étant capable d'expliquer pourquoi une telle décision a été prise (*explicabilité*). Pour cela, une série de questions est posée à l'utilisatrice jusqu'à avoir suffisamment d'informations sur ses préférences pour lui recommander la meilleure alternative. Plusieurs stratégies de questionnement peuvent être mises en place et l'une des plus populaires, l'élicitation incrémentale robuste, qui cherche à minimiser l'effort cognitif et posant le moins de questions possible (préférez-vous ce plan, ou celui-là?). Néanmoins, cette stratégie d'élicitation ne garantit pas une solution optimale quand l'utilisatrice est incertaine de ses préférences ou/et commet des erreurs; ou que les interactions entre les différents critères des alternatives ne sont pas modélisées correctement. Pire encore, cette stratégie peut conduire à une mauvaise recommandation, sans signal d'alerte.

Dans des précédents travaux [AD21], nous avons proposé une stratégie d'élicitation qui permet, grâce à la modélisation de l'incertitude des réponses via la théorie des possibilités, de détecter quand des réponses deviennent incohérentes entre-elles (soit car des réponses sont fausses, soit quand le modèle supposé n'est pas suffisamment riche pour prendre en compte des interactions complexes). Ainsi, l'élicitation est plus robuste face à des incohérences, permettant d'avoir des meilleures recommandations. De plus, nous avons proposé dans d'autres travaux [AD24] des stratégies qui permettent de restaurer la cohérence en présence de mauvaises réponses, permettant à la fois d'améliorer la qualité de la recommandation finale pour l'utilisatrice, mais aussi d'obtenir des informations précieuses sur quelles sont les possibles mauvaises réponses.

2 Objectifs

Les objectifs de ce projet sont multiples et font appels à différents domaines scientifiques.

Le premier objectif, déjà en partie avancé avec de premières solutions fonctionnelles qui restent à analyser, est de proposer une stratégie de restauration de cohérence en cas d'erreur de modèle. La solution actuellement étudiée est, en cas de détection d'incohérence, de garder les réponses précédemment obtenues, mais de changer pour un modèle de préférence qui soit



plus expressif et qui permet donc de capturer des interactions plus complexes entre les différents critères des alternatives (voir [Mar09] pour une présentation de différents modèles). Ce choix n'est pas anodin, car il faut opter pour un modèle suffisamment riche pour capturer les interactions, mais pas d'aller directement vers un modèle trop complexe, car avec l'augmentation du nombre de paramètres du modèle, les informations déjà obtenues via les réponses n'apporteront que peu d'information sur la recommandation à prendre.

Un deuxième objectif, plus ambitieux, est de proposer une stratégie de questionnement qui cherche à la fois à minimiser le nombre de questions en se focalisant sur une recommandation, mais qui explore aussi d'autres pistes pour détecter au plus vite de l'incohérence entre les réponses (dû possiblement au mauvais choix modèle), nécessitant un travail sur les problématiques d'exploration/exploitation, en trouvant un compromis entre convergence et détection d'incohérence. Par la suite, et en lien avec le ML, il pourrait être intéressant de déterminer la source de l'incohérence au travers des réponses obtenues : une utilisatrice qui donne involontairement de mauvaises réponses, ou un modèle de préférence inadéquat. Jusqu'à présent, nous proposons des corrections adaptées seulement à l'un des deux cas, en supposant au préalable quel était le cas. Déterminer la source d'incohérence nécessite d'analyser les réponses, qui restent peu riches en informations et limitées en nombre. Il est donc intéressant d'analyser la structure des réponses et de voir notamment si une stratégie basée sur l'exploration/exploitation aiderait la détection du type d'incohérence.

3 Candidature recherchée

Le ou la candidat(e) doit être en M2 et avoir des compétences en informatique et en programmation. Le stage se déroulera au LIAS, ISAE-ENSMA à Chasseneuil du Poitou, site du Futuroscope.

Le dossier de candidature doit être envoyé par courrier électronique, à loic.adam@ensma.fr et allel.hadjali@ensma.fr, et doit être constitué de :

- Lettre de motivation,
- CV mentionnant explicitement le parcours académique,
- Tout élément jugé utile pour mettre la candidature en valeur.

Nous recherchons en particulier :

- Connaissances en science des données (machine learning, statistiques, probabilités) et éventuellement en optimisation linéaire,
- Maîtrise des outils et des langages de programmation tels que Python ou Julia,
- Maîtrise du français et de l'anglais, tant à l'écrit qu'à l'oral.

Références

- [AD21] Loïc ADAM et Sébastien DESTERCKE. « Possibilistic preference elicitation by minimax regret ». In : *Uncertainty in artificial intelligence*. PMLR. 2021, p. 718-727.
- [AD24] Loïc ADAM et Sébastien DESTERCKE. « Handling inconsistency in (numerical) preferences using possibility theory ». In : *Information Fusion 103* (2024), p. 102089.
- [Mar09] Jean-Luc MARICHAL. « Aggregation functions for decision making ». In : *arXiv preprint arXiv :0901.4232* (2009).

