

Calcul de meilleures explications pour les forêts aléatoires : Raisons k -majoritaires et Raisons probabilistes majoritaires

Louenas Bounia*

*Université Sorbonne Paris Nord, CNRS, LIPN, 93430 Villetaneuse

Dans cet article, nous nous sommes intéressés au problème de l'explicabilité dans les forêts aléatoires, en introduisant des approches permettant de produire des explications à la fois robustes et concises : les raisons k -majoritaires et les raisons δ -probables majoritaires¹. Ces méthodes visent à répondre aux limitations cognitives des utilisateurs, tout en améliorant le niveau d'interprétabilité des modèles de forêts aléatoires.

Les raisons majoritaires MAJ (Audemard et al., 2022c) constituent des explications permettant de comprendre les décisions des forêts aléatoires en identifiant les ensembles minimaux de caractéristiques qui sont en accord avec la prédiction du modèle. Cependant, bien qu'elles soient efficaces, ces raisons peuvent manquer de robustesse ou être difficiles à interpréter en raison de leur grande taille. Pour surmonter ces limitations, nous avons étendu ce concept en introduisant les raisons k -MAJ, qui renforcent l'explication des prédictions de la forêt en exigeant qu'un sous-ensemble minimal soutienne non seulement la stricte majorité des arbres, mais aussi le plus grand nombre possible d'arbres de la forêt. Ces raisons (notées k -MAJ), y compris celles de taille minimale, améliorent la robustesse de l'explication en capturant les sous-ensembles qui valident les k arbres, tout en maintenant l'idée d'impliquer la majorité des arbres, qui est similaire à celle des MAJ mais en augmentant le nombre d'arbres impliqués, et en conservant un temps de calcul comparable. Les raisons k -MAJ constituent ainsi une version améliorée des raisons majoritaires.

Toutefois, malgré leurs avantages, les raisons k -MAJ et k -MAJ (même celles de taille minimale) peuvent encore être de grande taille pour être interprétables par les humains en raison de nos limitations cognitives (Miller, 1956). C'est pourquoi nous avons proposé les raisons δ -probables majoritaires, une approche probabiliste ajustable. Ces raisons permettent de contrôler la taille et l'intelligibilité des explications en paramétrant le seuil δ , qui définit la probabilité que les raisons identifiées garantissent la décision initiale des arbres de la forêt avec une probabilité élevée pour la majorité des arbres constituant la forêt. Cette flexibilité facilite l'adaptation des explications à différentes applications et permet aux utilisateurs de mieux les comprendre.

Nous avons mené des expérimentations approfondies pour évaluer l'utilité de ces approches. Nos résultats montrent que les raisons probabilistes majoritaires offrent une réduction significative de la taille des explications par rapport aux raisons directes (Izza et al., 2020; Audemard et al., 2022b,c), MAJ et même celles de taille minimale (notés minMAJ). Les boxplots présentés dans la figure 1 pour les ensembles de données « biomed » et « compas » montrent un extrait de nos résultats, illustrant la taille des explications avec des seuils de confiance de 70%, 80%

1. Une explication δ -probable majoritaire est une explication δ -probable (Louenas, 2024) pour la stricte majorité des arbres d'une forêt aléatoire.

Calcul de meilleures explications pour les forêts aléatoires

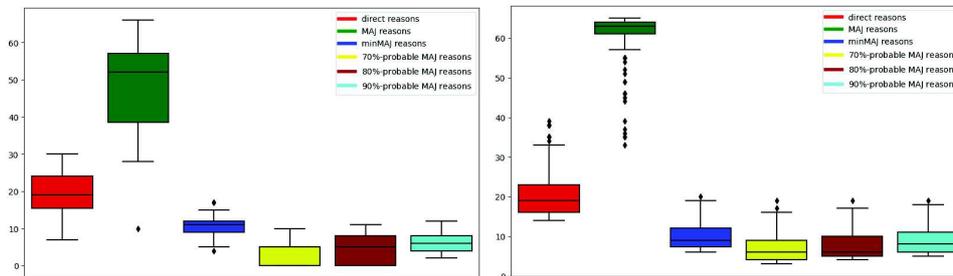


FIG. 1 – *Boxplots pour « biomed » et « compas », représentant les tailles des raisons directes, des raisons MAJ, des raisons minMAJ, et des raisons {70%, 80%, 90%}-probables majoritaires.*

et 90%. Ces expérimentations visent à mettre en évidence l'équilibre atteint par ces méthodes entre concision des explications et robustesse, avec des raisons adaptées aux capacités cognitives des utilisateurs.

Nos résultats expérimentaux, illustrés par les boxplots de la figure 1, mettent en évidence les avantages des approches proposées. Pour les ensembles de données « biomed » et « compas », les tailles des explications sont comparées entre les raisons directes, MAJ, minMAJ, et les raisons (δ -probables majoritaires) pour des seuils de confiance de 70%, 80% et 90%. Les raisons (δ -probables majoritaires) offrent une réduction significative de la taille, ce qui les rend particulièrement adaptées aux contextes où l'intelligibilité est primordiale. Ces résultats confirment la capacité des approches probabilistes à équilibrer concision et interprétabilité par rapport aux explications à base d'impliquants premiers (dites aussi explications abductives (Izza et Marques-Silva, 2021), ou raisons suffisantes (Audemard et al., 2022a)).

Références

- Audemard, G., S. Bellart, L. Bounia, F. Koriche, J. Lagniez, et P. Marquis (2022a). On preferred abductive explanations for decision trees and random forests. In *Proc. of IJCAI'22*.
- Audemard, G., S. Bellart, L. Bounia, F. Koriche, J.-M. Lagniez, et P. Marquis (2022b). On the explanatory power of boolean decision trees. *Data & Knowledge Engineering* 142, 102088, doi: <https://doi.org/10.1016/j.datak.2022.102088>.
- Audemard, G., S. Bellart, L. Bounia, F. Koriche, J.-M. Lagniez, et P. Marquis (2022c). Trading complexity for sparsity in random forest explanations. In *Proc. of AAAI'22*.
- Izza, Y., A. Ignatiev, et J. Marques-Silva (2020). On explaining decision trees. *ArXiv abs/2010.11034*.
- Izza, Y. et J. Marques-Silva (2021). On explaining random forests with SAT. In *Proc. of IJCAI'21*, pp. 2584–2591.
- Louenas, B. (2024). Enhancing the Intelligibility of Boolean Decision Trees with Concise and Reliable Probabilistic Explanations. In *20th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, Lisboa, Portugal.
- Miller, G. A. (1956). The magical number seven, plus or minus two : Some limits on our capacity for processing information. *The Psychological Review* 63(2), 81–97.