Projet SDP

NATHAN BRUCKMANN

JÉRÔME AUGUSTE

RUBEN PARTOUCHE

Sommaire

Modèles

MR-SORT

NCS MONOTONE

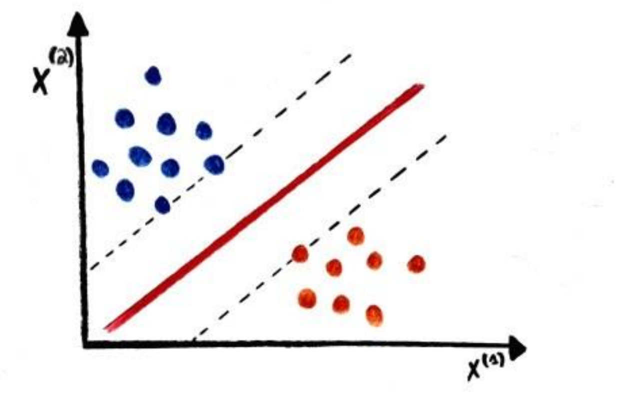
NCS NON MONOTONE

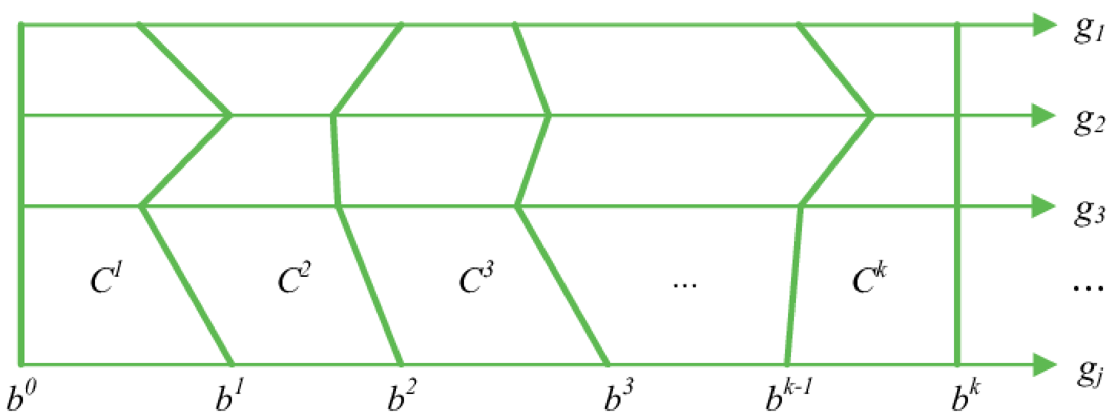
MAX SAT

MR-Sort

Méthode par programmation linéaire.
Objectif : trouver les frontières, le poids de chaque critère, et la limite lambda.

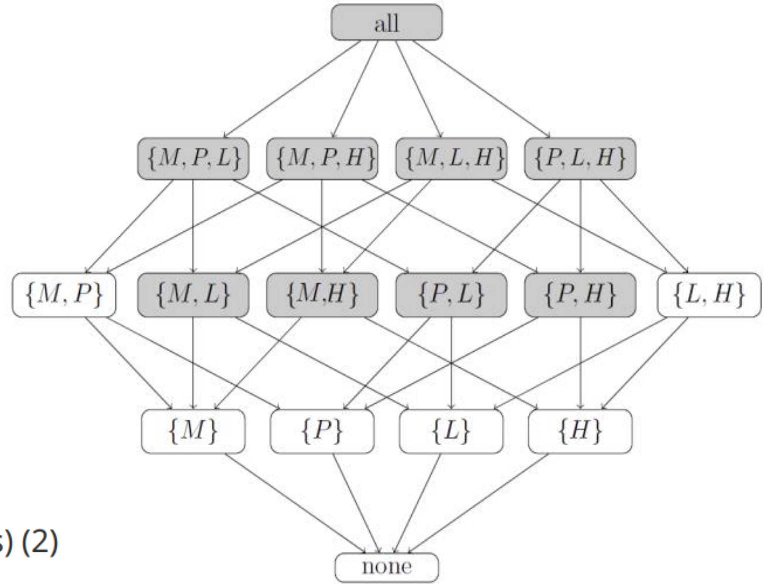
$$
\begin{cases}
max\ \alpha & \\
\quad \sum_{i \in N} c_{ij} + x_j + \varepsilon = \lambda & \forall a_j \in A^{*1} \\
\quad \sum_{i \in N} c_{ij} = \lambda + y_j & \forall a_j \in A^{*2} \\
\quad \alpha \leq x_j,\ \alpha \leq y_j & \forall a_j \in A^{*} \\
\quad c_{ij} \leq w_i & \forall a_j \in A^{*}, \forall i \in N \\
\quad c_{ij} \leq \delta_{ij} & \forall a_j \in A^{*}, \forall i \in N \\
\quad c_{ij} \geq \delta_{ij} - 1 + w_i & \forall a_j \in A^{*}, \forall i \in N \\
\quad M\delta_{ij} + \varepsilon \geq g_i(a_j) - b_i & \forall a_j \in A^{*},\ \forall i \in N \\
\quad M(\delta_{ij} - 1) \leq g_i(a_j) - b_i & \forall a_j \in A^{*},\ \forall i \in N \\
\quad \sum_{i \in N} w_i = 1,\ \lambda \in [0.5, 1] & \\
\quad w_i \in [0, 1] & \forall i \in N \\
\quad c_{ij} \in [0, 1],\ \delta_{ij} \in \{0, 1\} & \forall a_j \in A^{*},\ \forall i \in N \\
\quad x_j, y_j \in \mathbb{R} & \forall a_j \in A^{*} \\
\quad \alpha \in \mathbb{R} &
\end{cases}
$$





NCS

Méthode par solveur de satisfiabilité
Objectif : trouver les parties représentant une **coalition majoritaire** parmi l'ensemble des parties de critères **monotones.**

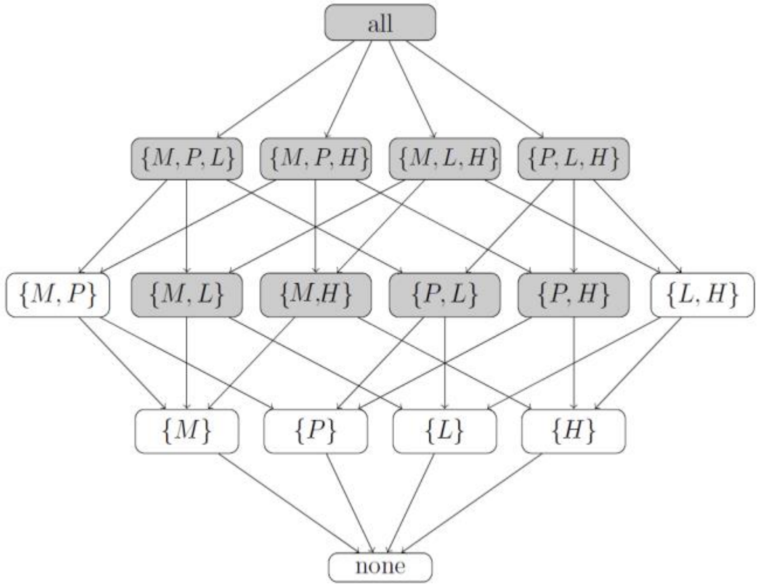


- $x_{i,h,k}$ for each criterion $i \in \mathcal{N}$, for each boundary $1 \leq h \leq p-1$ and each grade k of that criterion assuming the extreme classes undefined boundary are the extreme values taken by the grades
- y_B for each coalition $B \in \mathcal{P}(\mathcal{N})$

- $\forall i \in \mathcal{N}, \forall 1 \leq h \leq p-1, \forall k < k', \quad x_{i,h,k} \Rightarrow x_{i,h,k'}$ (We can only consider adjacent grades) (1)
- $\forall i \in \mathcal{N}, \forall 1 \leq h < h' \leq p-1, \forall k, \quad x_{i,h',k} \Rightarrow x_{i,h,k}$ (we can only consider adjacent boundaries) (2)
- $\forall B \subset B' \subseteq \mathcal{N}, \quad y_B \Rightarrow y_{B'}$ (We can only consider B and B' such that $|B' \setminus B| = 1$) (3)
- $\forall B \subseteq \mathcal{N}, \forall 1 \leq h \leq p-1 \forall u \in X^* : A(u) = C^{h-1}, \quad \bigwedge_{i \in B} x_{i,h,u_i} \Rightarrow \neg y_B$ (4)
- $\forall B \subseteq \mathcal{N}, \forall 1 \leq h \leq p-1 \forall a \in X^* : A(a) = C^{h}, \quad \bigwedge_{i \in B} \neg x_{i,h,a_i} \Rightarrow y_{\mathcal{N} \setminus B}$ (5)

NCS Single Peak

Méthode par solveur de satisfiabilité
Objectif : trouver les parties représentant une **coalition majoritaire** parmi l'ensemble des parties de critères **monotones et non monotones (Single Peak).**



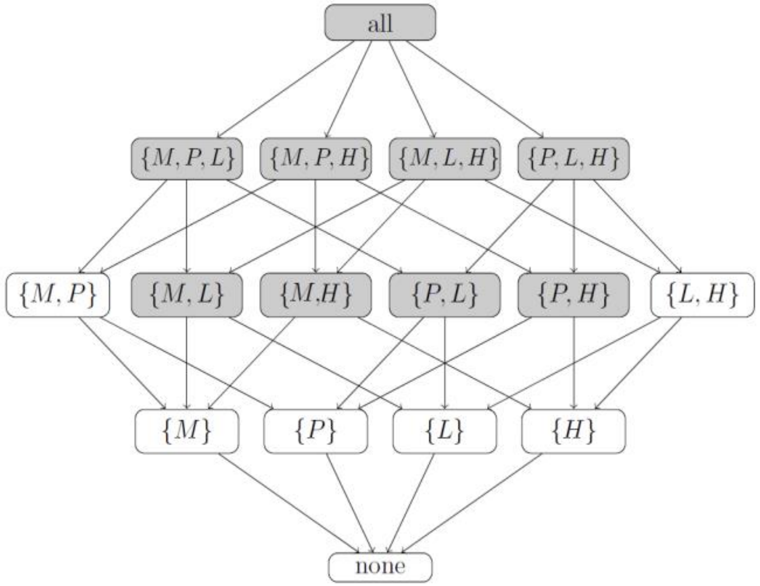
On remplace simplement la clause (1) de monotonie par :

- $\forall i \in \mathcal{N}, \forall 1 \leq h \leq p-1, \forall k < k' < k'', \quad x_{i,h,k} \wedge x_{i,h,k''} \Rightarrow x_{i,h,k'}$ (1')

NCS Max-SAT

Méthode par solveur de satisfiabilité

Objectif : trouver les parties représentant une **coalition majoritaire** parmi l'ensemble des parties de critères **monotones et non monotones**, en minimisant le nombre de clauses non satisfaites parmi les clauses issues de l'ensemble d'entrainement lorsque le problème n'a pas de solution



Implémentation de la génération de données

CAS MONOTONE

CAS NON MONOTONE

DÉSÉQUILIBRE DE CLASSES

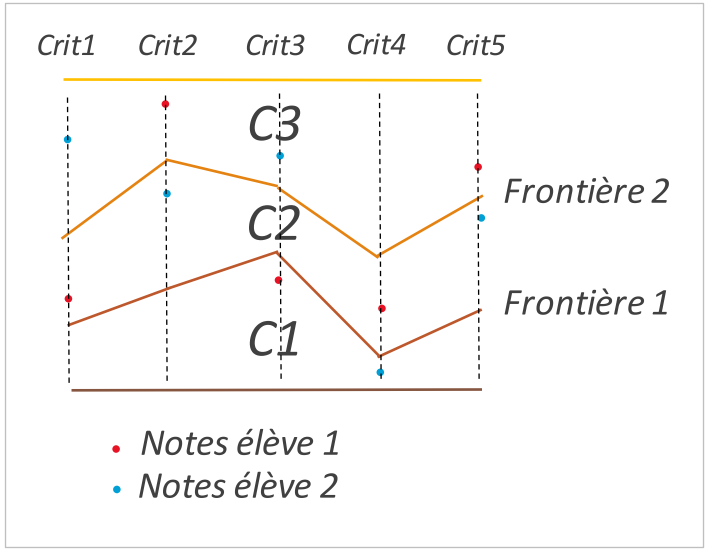
Implémentation et générateur : Cas monotone

- On tire **lambda** dans une loi uniforme (0.5,1)

- On tire un **vecteur de poids** dans une **loi normale** centrée en 2 (pour éviter les poids négatifs), puis on normalise le vecteur

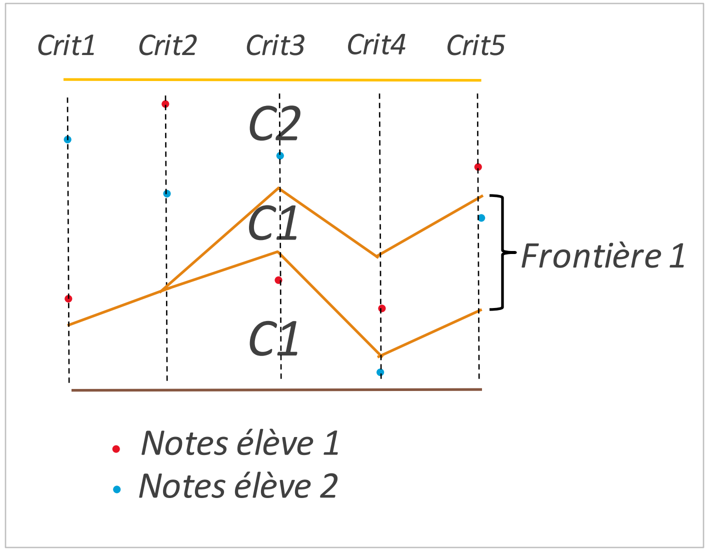
- On tire les **frontières** pour chaque critère selon une **loi uniforme** entre la frontière précédente et une limite supérieure fixée en fonction du critère, de manière à assurer la dominance de chaque frontière sur la précédente.

- Connaissant tous les paramètres, on affecte chaque élève dans une catégorie par calcul direct.



Implémentation et générateur : Cas non monotone

- On tire les **frontières** pour chaque critère dans une **loi uniforme** entre les bornes de la frontière précédente, en triant la liste a deux éléments ainsi obtenus



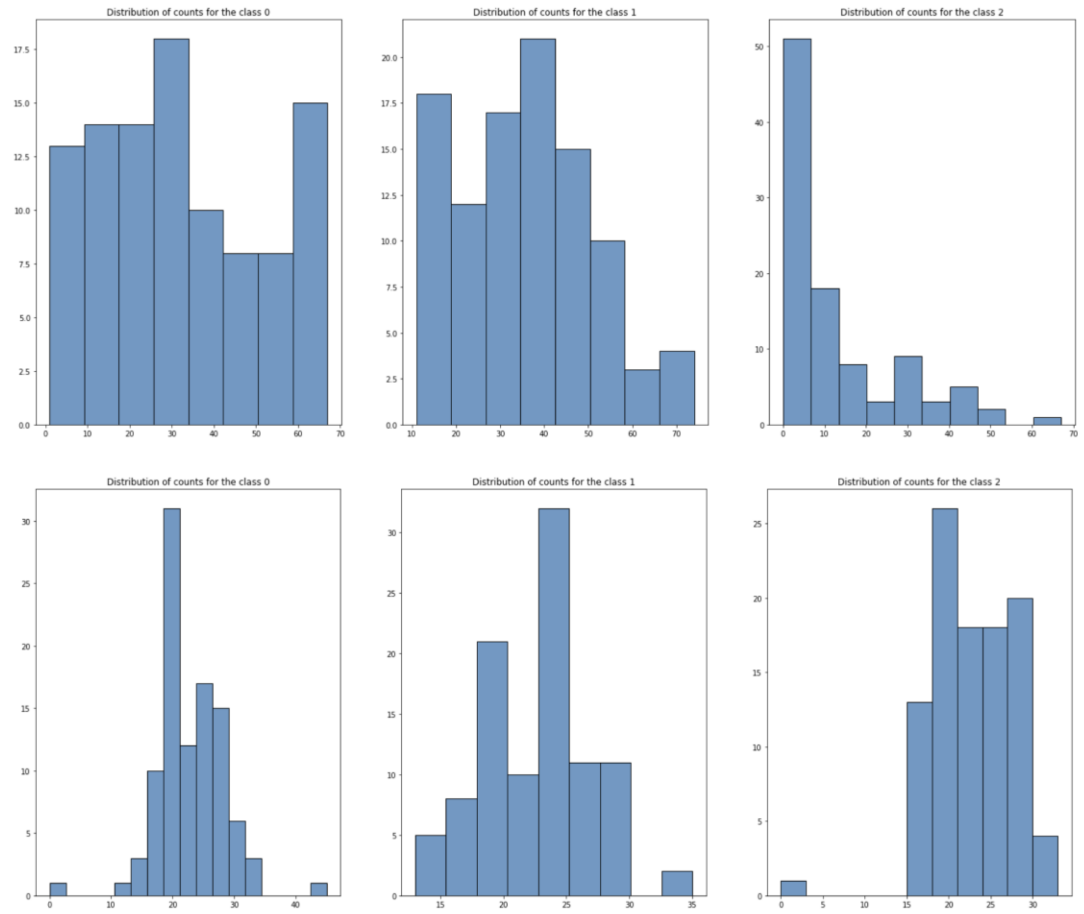
Implémentation et générateur : Déséquilibre de classe

On obtient parfois des datasets très déséquilibrés en procédant ainsi, avec des conséquences négatives :

- Mauvais apprentissage
- Modèle biaisé
- Performances finales biaisées

Pour éviter cela, on resample les notes avec une technique d'augmentation de données tant que toutes les classes ne sont pas présentes.

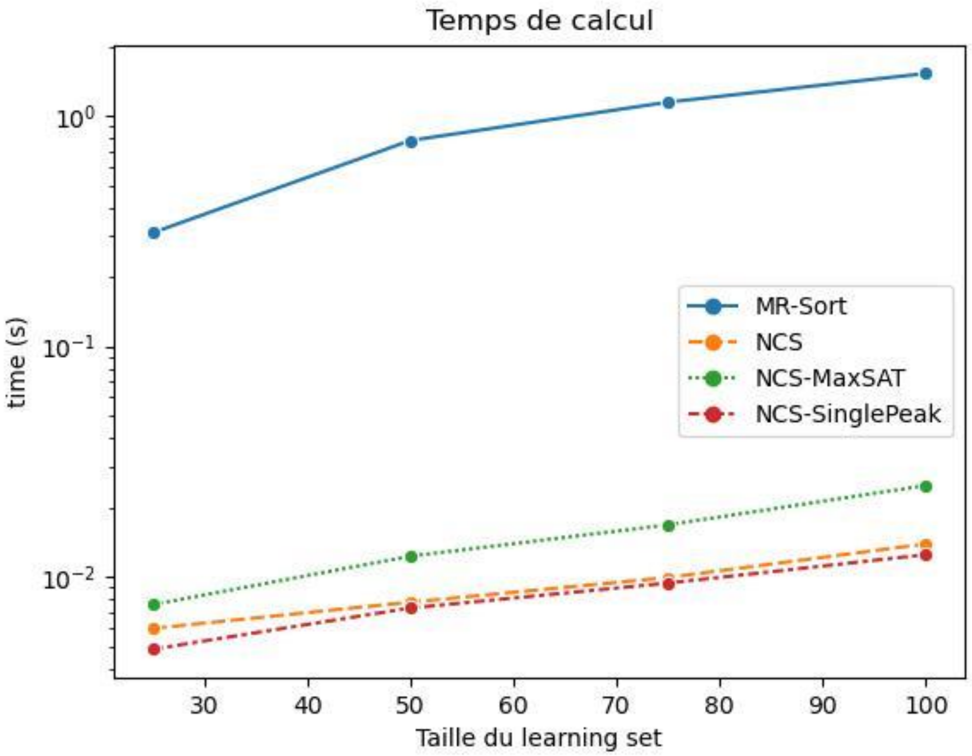
On évite de tirer trop longtemps en fixant un nombre d'itération maximal, et en appliquant un coefficient dégressif sur le paramètre qui fait décroitre l'écart maximal admissible avec la distribution parfaitement équilibrée.

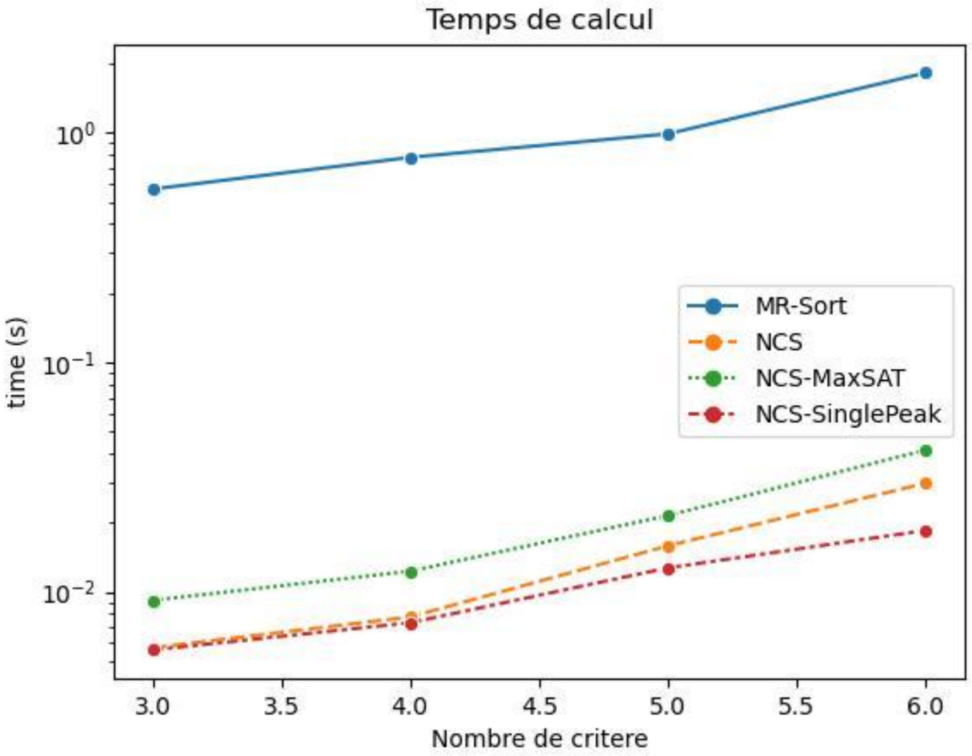


Résultats

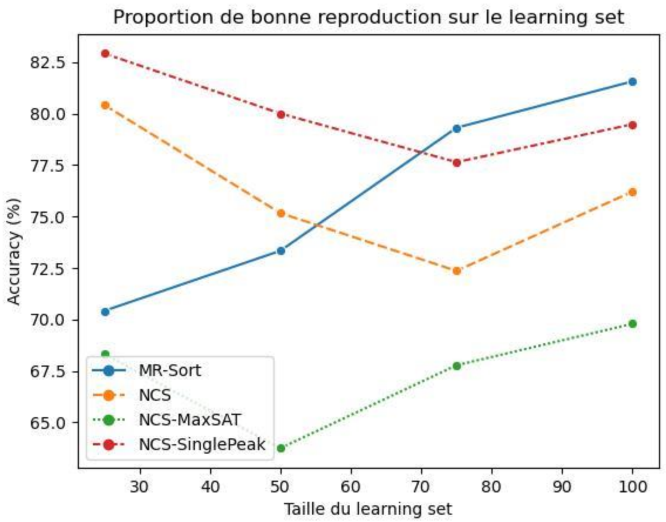
- TEMPS DE CALCUL
- ACCURACY
- BRUIT
- NOMBRE DE CRITÈRES ET DE CLASSES

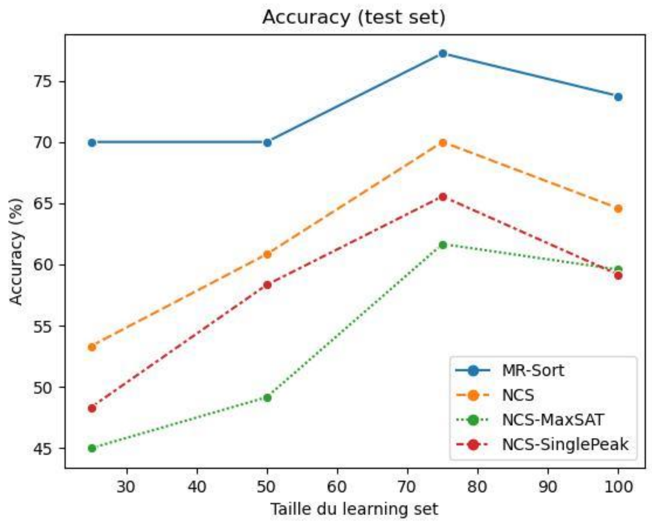
Résultats : Temps de calcul

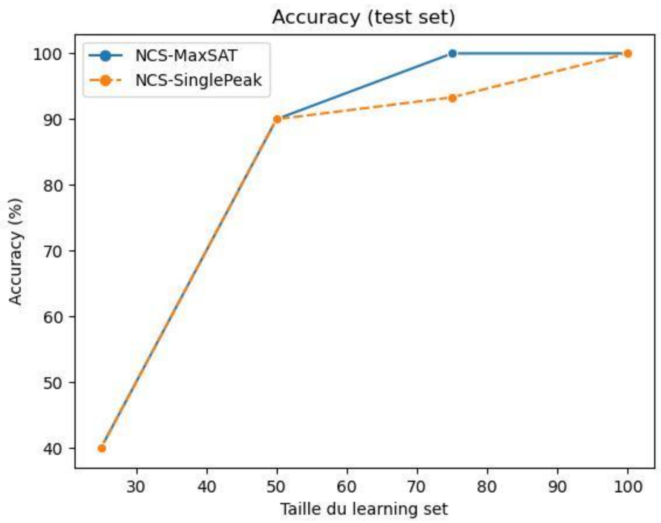




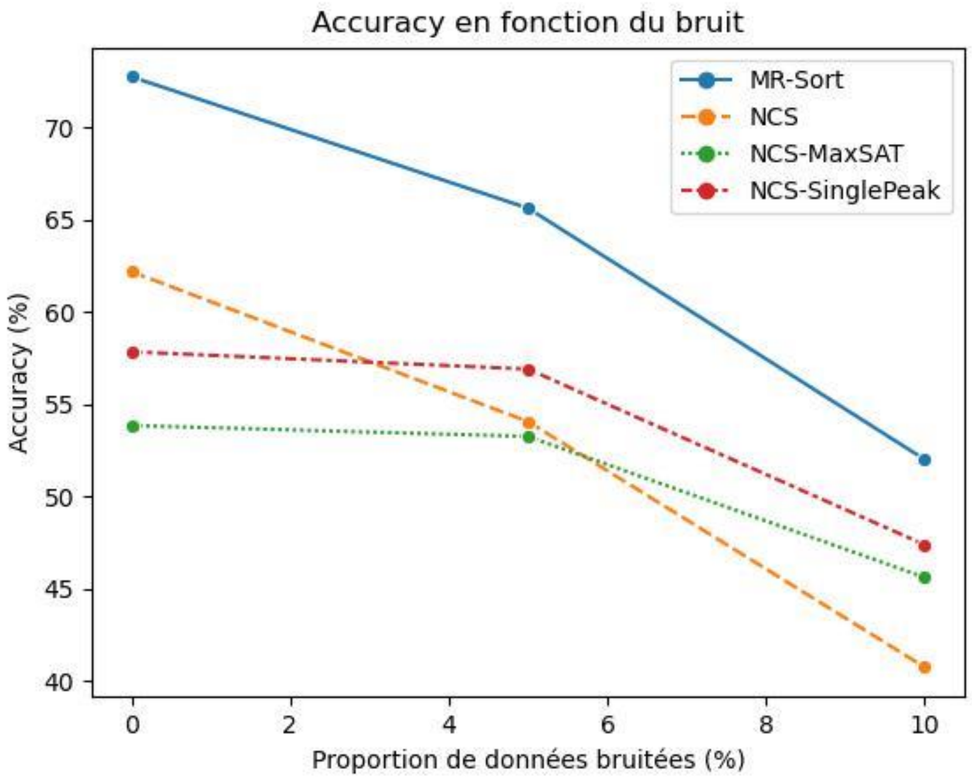
Résultats : Accuracy



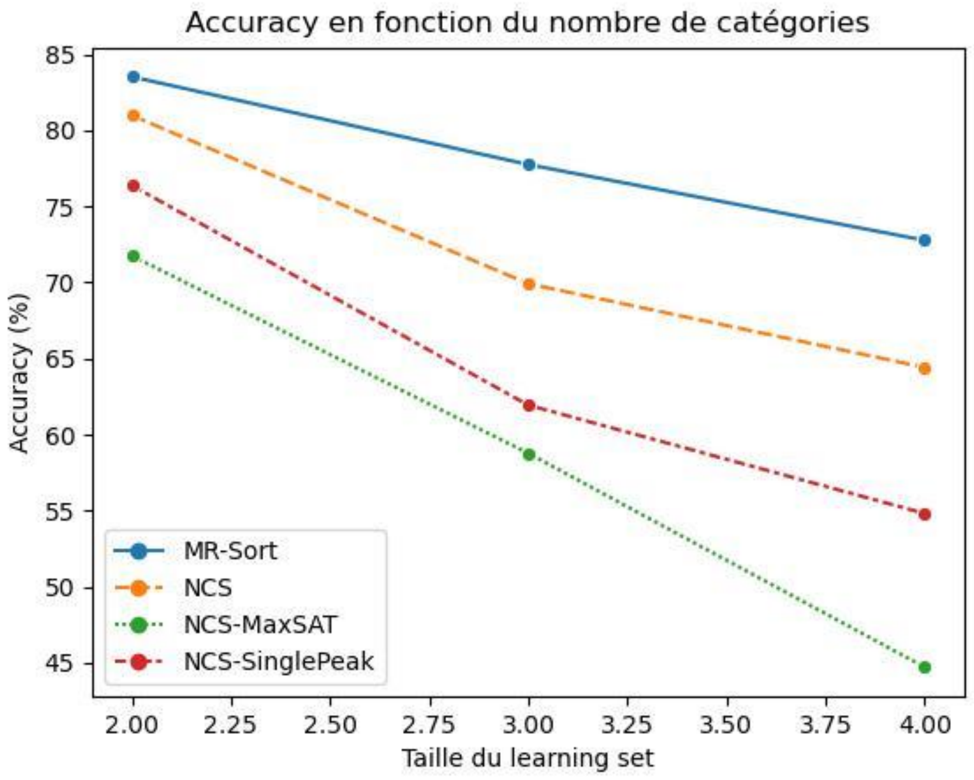


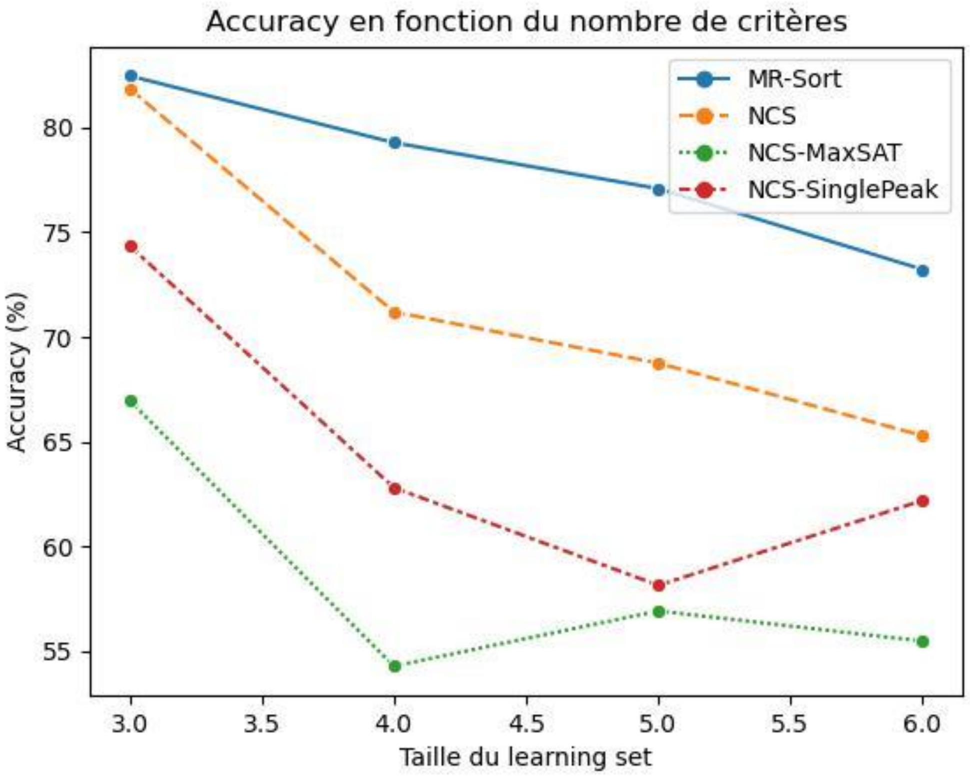


Résultats : Bruit



Résultats : Nombre de critères et de classes





Conclusion

MR-SORT

- Bonne précision sur le problème monotone
- Assez robuste face au bruit
- Donne toujours un résultat
- Temps de calcul

SAT

- Se transforme facilement pour adapter le problème
- Temps de calcul
- Précision satisfaisante selon les conditions expérimentales
- Ne donne pas toujours de résultat sauf en formulation maxSAT
- Sensible au bruit

Pour aller plus loin : tester les problèmes single-peak et maxsat sous la forme de programmes d'optimisation linéaire, pour comparer les performances en terme de précision et temps de calcul (probablement fortement dégradé)

Sources

K. Belahcène, C. Labreuche, N. Maudet, V. Mousseau, W. Ouerdane, *An efficient SAT formulation for learning multiple criteria non-compensatory sorting rules from examples*, Computers & Operations Research, Volume 97, 2018, Pages 58-71

Ali Tlili, Khaled Belahcène, Oumaima Khaled, Vincent Mousseau, Wassila Ouerdane, *Learning non-compensatory sorting models using efficient SAT/MaxSAT formulations*, European Journal of Operational Research, Volume 298, Issue 3, 2022, Pages 979-1006

Leroy, Agnès & Mousseau, Vincent & Pirlot, Marc. (2011). *Learning the Parameters of a Multiple Criteria Sorting Method*. 219-233. 10.1007/978-3-642-24873-3_17.