



HAL
open science

Apprentissage de contraintes : résumé et application sur un problème de planification

David Tremblet, Simon Thevenin, A. Dolgui

► To cite this version:

David Tremblet, Simon Thevenin, A. Dolgui. Apprentissage de contraintes : résumé et application sur un problème de planification. ROADEF 2023 : 24ème congrès annuel de la Société Française de Recherche Opérationnelle et d'Aide à la Décision, Feb 2023, Rennes, France. hal-04074975

HAL Id: hal-04074975

<https://hal.science/hal-04074975>

Submitted on 19 Apr 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Apprentissage de contraintes : résumé et application sur un problème de planification

David Tremblet¹, Simon Thevenin¹, Alexandre Dolgui¹

IMT Atlantique, LS2N-CNRS, La Chantrerie, 4 Rue Alfred Kastler, B.P. 20722, 44307 Nantes, France

david.tremblet@imt-atlantique.fr

Mots-clés : *Machine Learning, Data-Driven methods, Planification et Ordonnancement*

1 Contexte

Les méthodes de machine learning ont suscité un véritable engouement dans la communauté de recherche opérationnelle, et elles se révèlent très efficaces sur des problèmes d'optimisation basée sur l'analyse prédictive de données. Ces méthodes d'apprentissages trouvent leur intérêt dans le fait de mieux capturer certaines tendances observables au travers de jeux de données, offrant de nouvelles perspectives de résolution lorsqu'elles sont appliquées à des problèmes d'optimisation. Un exemple concret est l'utilisation de méthodes basées sur de la régression pour approcher un paramètre incertain tel que des temps de trajet d'un problème de transport ou la demande d'un problème de planification. D'autres approches consistent en l'utilisation de machine learning pour aiguiller des solveurs de programmation mathématique dans leur prise de décisions, préservant ainsi un temps de calcul précieux lors de la résolution d'un problème (BENGIO, LODI et PROUVOST 2021).

La branche de l'apprentissage machine qui nous intéresse ici concerne la traduction de ces méthodes de machine learning en modèles mathématiques, afin de remplacer des contraintes mathématiques difficiles à modéliser. Si l'approche peut paraître triviale pour régression linéaires classique, elle nécessite davantage d'effort pour traduire d'autres modèles basée sur des arbres de décision ou des réseaux de neurones, et l'utilisation de ces méthodes une fois traduit a suscité un intérêt grandissant dans la communauté de recherche opérationnelle ces dernières années (MIŠIĆ 2020 ; FISCHETTI et JO 2018).

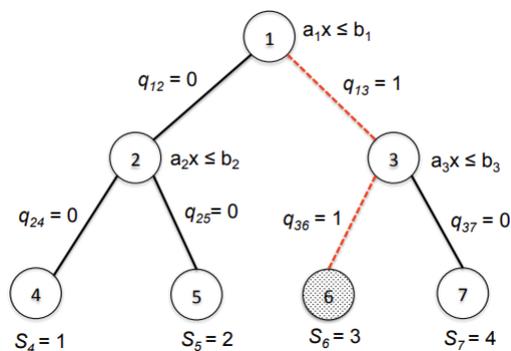


FIG. 1 – Traduction d'un arbre de régression en équations mathématiques (BIGGS et HARISS 2017)

La figure 1 illustre un exemple d'apprentissage de contrainte pour un modèle basé sur un arbre de décision. Chaque noeud de l'arbre représente une requête, dont la réponse dépend de la valeur prise par les variables de décision x . Une fois intégrées à un modèle mathématiques, ces équations vont inférer, par le biais de variables binaires q , une prédiction S qui servira de substitut à une ou plusieurs équations à remplacer dans le modèle mathématique. Dans cette présentation, nous expliquerons en détail comment intégrer ces modèles d'apprentissage machine dans un programme linéaire en nombres entiers, et nous présenterons les avantages de ce type d'approche pour l'optimisation. Pour illustrer nos propos, nous nous pencherons sur un problème de planification et ordonnancement.

2 Description de la méthode et application

Nous considérons dans cette étude un problème de détermination de lots (ou *lot-sizing*) et ordonnancement intégré (DAUZÈRE-PÉRES et LASSERRE 1994). Dans ce type de problème, l'élaboration d'un plan de production optimal doit généralement prendre en compte une consommation de capacité à chaque période de l'horizon de production, qui se définit le plus souvent par la résolution d'un problème d'ordonnancement. Cette étape garantit que les quantités à produire dans l'atelier sont bien réalisables, et permet d'éviter un cas de production tendu qui engendrerait des retards sur la production, entraînant des coûts supplémentaires non négligeables. Or, la considération simultanée de la planification et de l'ordonnancement constitue un véritable challenge, puisqu'elle nécessite de résoudre en parallèle de deux problèmes difficiles. Dans notre exemple, le premier problème est connu comme étant NP-Difficile, car des contraintes de capacité sont en jeu, alors que le deuxième est connu comme étant NP-Complet dans notre cas, puisque l'on considère un problème de faisabilité.

L'idée de notre approche serait d'approcher les solutions trouvées par l'ordonnancement en générant un jeu de données massif qui sera ensuite soumis à un ou plusieurs modèles d'apprentissage automatique. Une fois les données du problème acquises, le modèle de machine learning est alors entraîné puis traduit sous la forme d'un programme linéaire en nombre entier (MARAGNO et al. 2021). Un tel modèle sera capable de fournir une bonne approximation du makespan du problème d'ordonnancement en question et il sera alors possible de substituer les contraintes liées à l'ordonnancement dans notre problème de planification (DIAS et IERAPETRITOU 2019).

Les modèles concernés par nos travaux sont la régression linéaire, du fait de leur simplicité de modélisation, ainsi que les arbres de régression, qui restent facilement interprétable et permettent de capturer des tendances non linéaires d'un jeu de données. Une autre méthode envisagée se base sur un modèle d'arbres de décision et régression linéaire combinés (ou *Model Trees*, QUINLAN 1992), qui offre une meilleure approximation dans notre cas.

Au choix de la méthode s'ajoute une difficulté d'apprentissage pour notre cas, puisqu'il convient de garantir que la date de fin du plan de production ne soit pas sous-estimée. Pour palier à ce problème, nous étudierons sur une méthode d'apprentissage spécifique, ainsi qu'une méthode de génération d'exemples contradictoires.

Remerciements

Ce travail a été réalisé dans le cadre du projet ASSISTANT financé par la Commission européenne, sous le numéro de subvention 101000165, H2020 - ICT-38-2020. Les auteurs tiennent également à remercier la région Pays de la Loire pour son soutien financier.

Références

- BENGIO, Yoshua, Andrea LODI et Antoine PROUVOST (avr. 2021). « Machine learning for combinatorial optimization : A methodological tour d'horizon ». In : *European Journal of Operational Research*.
- BIGGS, Max et Rim HARISS (2017). « Optimizing Objective Functions Determined from Random Forests ». In : *SSRN Electronic Journal*. DOI : 10.2139/ssrn.2986630.
- DAUZÈRE-PÉRES, Stephane et Jean-Bernard LASSERRE (juin 1994). « Integration of lotsizing and scheduling decisions in a job-shop ». In : *European Journal of Operational Research*.
- DIAS, Lisia S. et Marianthi G. IERAPETRITOU (août 2019). « Data-driven feasibility analysis for the integration of planning and scheduling problems ». In : *Optimization and Engineering*.
- FISCHETTI, Matteo et Jason JO (2018). « Deep neural networks and mixed integer linear optimization ». In : *Constraints*.
- MARAGNO, Donato et al. (2021). « Mixed-Integer Optimization with Constraint Learning ». In : arXiv : 2111.04469 [math.OC].
- MIŠIĆ, Velibor V. (sept. 2020). « Optimization of Tree Ensembles ». In : *Operations Research*.
- QUINLAN, John R (1992). « Learning with continuous classes ». In : *5th Australian joint conference on artificial intelligence*. World Scientific.