

Introduction

Rachid CHELOUAH

CY Cergy Paris Université, Cergy-Pontoise, France

L'apprentissage automatique révolutionne notre monde. Il est difficile de concevoir une autre technologie de l'information qui se soit développée aussi rapidement ces dernières années, en matière d'impact réel.

Les domaines de l'apprentissage automatique et de l'optimisation sont fortement imbriqués. En effet, les problèmes d'optimisation sont au cœur des méthodes d'apprentissage automatique et les algorithmes d'optimisation modernes utilisent de plus en plus l'apprentissage automatique pour améliorer leur efficacité.

L'apprentissage automatique a des applications dans tous les domaines de la science. Il existe de nombreuses méthodes d'apprentissage, chacune d'entre elles utilisant une structure algorithmique différente pour optimiser les prédictions, en fonction des données reçues. Le premier objectif de cet ouvrage est donc de mettre en lumière les principes et méthodes clés qui sont communs aux deux domaines.

L'apprentissage automatique et l'optimisation partagent trois composantes : la représentation, l'évaluation et la recherche itérative. Pourtant, alors que les solveurs d'optimisation sont généralement conçus pour être rapides et précis sur des modèles implicites, les méthodes d'apprentissage automatique doivent être génériques et entraînées hors ligne sur des ensembles de données. Les problèmes d'apprentissage automatique présentent de nouveaux défis pour les chercheurs en optimisation ; et les praticiens de l'apprentissage automatique recherchent des algorithmes d'optimisation plus simples et génériques.

Optimisation et apprentissage,

coordonné par Rachid CHELOUAH et Patrick SIARRY. © ISTE Editions 2023.

Assez récemment, les approches modernes de l'apprentissage automatique ont également été appliquées à la conception des algorithmes d'optimisation eux-mêmes, en tirant parti de leur capacité à capturer des informations utiles à partir de structures complexes dans de grands espaces. Ces capacités semblent utiles, notamment pour les composantes de représentation et d'évaluation. Comme les grandes structures complexes sont omniprésentes dans les problèmes d'optimisation et peuvent être utilisées comme d'énormes ensembles de données implicites, l'utilisation de l'apprentissage automatique a permis d'améliorer l'efficacité et la généralité des méthodes d'optimisation.

Cet ouvrage présente les avancées modernes en matière de sélection, de configuration et d'ingénierie des algorithmes qui reposent sur l'apprentissage automatique et l'optimisation. Il est structuré en deux parties. La partie 1 est consacrée aux applications d'optimisation les plus courantes. La partie 2 décrit et met en œuvre plusieurs applications de l'apprentissage automatique.

La partie 1 comprend quatre chapitres qui se concentrent sur l'application réelle des algorithmes d'optimisation.

Le chapitre 1 traite du problème de l'acheminement des véhicules avec des contraintes de chargement et combine deux problèmes d'optimisation combinatoire : le problème de tournées des véhicules avec contraintes de capacité (CVRP, *Capacitated Vehicle Routing Problems*) et le problème de mise en boîte à deux/trois dimensions (2/3D-BPP). Les auteurs ont étudié des problèmes de transport réels tels que le transport de meubles ou de machines industrielles.

L'objectif principal du chapitre 2 est de créer la solution d'ordonnancement la plus appropriée qui optimise plusieurs mesures de QoS simultanément. Les auteurs adaptent donc la métaheuristique largement utilisée, « l'algorithme génétique », comme méthode d'optimisation. L'approche d'ordonnancement proposée est testée en simulant une application IoT de soins de santé, modélisée comme un flux de travail, et plusieurs repères de flux de travail scientifiques. Les résultats montrent l'efficacité de l'approche proposée ; elle génère un plan d'ordonnancement qui optimise mieux les différentes métriques de QoS considérées.

Le chapitre 3 se concentre sur l'optimisation du loup gris (GWO) et son adaptation à un espace de recherche continu. Il commence par aborder la modélisation mathématique de l'optimisation dans un espace de recherche discret binaire. Des modules de binarisation sont ensuite fournis, permettant à des métaheuristicques continues de résoudre des problèmes de sélection de caractéristiques dans un espace de recherche binaire. Ces modules de binarisation sont ensuite utilisés pour créer la métaheuristique binaire bGWO. Enfin, une démonstration expérimentale montre les performances de bGWO dans la résolution

de problèmes de sélection de caractéristiques sur 18 jeux de données provenant du référentiel d'apprentissage machine de l'UCI.

L'objectif principal du chapitre 4 est d'aborder le problème d'équilibrage de la chaîne de montage à modèle mixte de type 2 avec des temps de tâche déterministes. Pour résoudre ce problème, une amélioration de la procédure de recherche adaptative aléatoire gloutonne, connue sous le nom de procédure de recherche adaptative aléatoire gloutonne réactive, est proposée. Cette version réactive est basée sur la variation de la valeur du paramètre de la liste restreinte de candidats, α . La GRASP réactive proposée est hybridée avec l'heuristique de poids positionnel classé pour construire les solutions initiales. Les résultats obtenus par la GRASP réactive hybride proposée sont comparés à ceux obtenus par la GRASP de base, démontrant l'effet du mécanisme d'apprentissage.

La partie 2 comprend quatre chapitres consacrés à l'intelligence artificielle, à l'apprentissage automatique et à leurs applications.

Le principal défi des systèmes de recommandation provient de la modélisation de la dépendance entre les différentes entités, en incorporant des informations multiformes telles que les préférences des utilisateurs, les attributs des articles et l'influence mutuelle des utilisateurs, ce qui donne lieu à des caractéristiques plus complexes. Pour faire face à ce problème, les auteurs du chapitre 5 conçoivent des modèles d'apprentissage automatique d'ensemble empilés pour les recommandations. Leur système de recommandation intègre un module de filtrage collaboratif (CF, *Collaborative Filtering*) et un module de recommandation par empilement. Un mécanisme d'attention interactif est ensuite introduit pour modéliser la relation d'influence mutuelle entre les utilisateurs et les articles. Des expériences sur des ensembles de données du monde réel démontrent que l'algorithme proposé permet d'obtenir des prédictions plus précises et une meilleure efficacité de recommandation.

Dans le domaine de l'audit interne, la capacité de traiter toutes les informations disponibles relatives à l'univers de l'audit ou au sujet pourrait améliorer la qualité des résultats. La classification des documents textuels d'audit (données non structurées) pourrait permettre d'utiliser des informations supplémentaires pour améliorer les données structurées existantes, créant ainsi un meilleur support de connaissances pour le processus d'audit. Le chapitre 6 présente une comparaison des résultats des algorithmes classiques d'apprentissage automatique et d'apprentissage profond, combinés à des incorporations de mots avancées pour classer les conclusions des rapports d'audit interne.

La conception d'une architecture de contrôle est un problème central dans un projet de réalisation d'un robot mobile autonome. En l'absence d'une solution générique, il est essentiel de proposer une nouvelle approche détaillant le processus de conception d'un

système intelligent capable de s'adapter à tous les changements de l'environnement de navigation. Le chapitre 7 propose d'utiliser le paradigme multi-agent et la logique floue dans la conception de l'architecture de contrôle pour la navigation autonome du robot mobile dans un environnement contraint. L'architecture de contrôle est conçue pour résoudre différents problèmes créés lors de la navigation. Elle est composée de quatre agents : l'agent de perception, l'agent de faisabilité, l'agent de locomotion et l'agent de contrôle flou.

La détection d'intrusion est un concept clé de la sécurité des réseaux informatiques modernes. Elle vise à analyser l'état actuel d'un réseau en temps réel et à identifier les anomalies potentielles du système, en les signalant dès qu'elles sont identifiées. Cela permet de détecter des logiciels malveillants jusqu'alors inconnus. Les réseaux neuronaux artificiels sont des algorithmes d'apprentissage automatique supervisés inspirés du cerveau humain. Ce type de réseau est un choix populaire parmi les techniques d'exploration de données d'aujourd'hui et s'est déjà avéré être un choix précieux pour la détection des intrusions. Dans le chapitre 8, l'auteur construit un réseau neuronal de type *feed-forward* entraîné sur le jeu de données NSL-KDD, afin de classer les connexions du réseau comme appartenant à l'une des deux catégories possibles : normale ou anormale. Son objectif est de maximiser le niveau de précision de la reconnaissance de nouveaux échantillons de données.