UNIVERSITÉ D'ARTOIS

Avis de Soutenance

Monsieur Mehdi SABIRI

Informatique et applications

Soutiendra publiquement ses travaux de thèse intitulés
Exploiter données et connaissances, des prédictions à leurs explications

Travaux dirigés par Monsieur Pierre MARQUIS et Monsieur Gilles AUDEMARD

Soutenance prévue le ***lundi 09 novembre 2026*** à 9h00
Lieu : 13 Rue Jean Souvraz, 62300 Lens
Salle : prestige

Composition du jury proposé

M. Pierre MARQUIS	Université d'Artois	Directeur de thèse
M. Sébastien DESTERCKE	Université de Technologie de Compiègne	Rapporteur
M. Laurent SIMON	Université de Bordeaux	Rapporteur
Mme Wassila OUERDANE	Centrale Supélec, Université Paris-Saclay	Examinatrice
M. Gilles AUDEMARD	Université d'Artois	Co-directeur de thèse
Mme Sylvie COSTE-MARQUIS	Université d'Artois	Examinatrice

Résumé :

L'intelligence artificielle explicable (eXplainable Artificial Intelligence, XAI) vise à rendre les modèles d'apprentissage automatique plus compréhensibles en fournissant des explications sur leurs prédictions, tout en permettant, dans certains cas, de corriger leurs erreurs. Cependant, un défi majeur subsiste : les modèles les plus performants sont généralement les plus opaques, tandis que les modèles les plus interprétables offrent souvent des performances prédictives plus limitées. Afin de réduire cette tension entre performance et explicabilité, les approches d'intelligence artificielle hybride proposent de combiner les méthodes statistiques de l'apprentissage automatique avec des connaissances symboliques exprimées sous forme de règles logiques ou de contraintes. Cette hybridation permet non seulement d'améliorer l'explicabilité des prédictions, mais également d'intégrer les connaissances et les préférences des utilisateurs afin de renforcer leur confiance dans les systèmes d'intelligence artificielle. Cette thèse s'inscrit dans ce contexte et vise à concilier performance prédictive et interprétabilité des modèles d'apprentissage automatique. Pour atteindre cet objectif, nous proposons un cadre hybride combinant apprentissage statistique, raisonnement logique et interaction humaine. Nos travaux portent sur l'extraction d'explications abductives pour les modèles à base d'arbres, la distillation incrémentale de modèles complexes vers des modèles plus interprétables, l'exploitation de connaissances extraites automatiquement à partir des données ainsi que l'intégration d'un expert humain dans un processus de correction et d'enrichissement des connaissances. Les méthodes proposées permettent d'améliorer simultanément la qualité des explications, la capacité de correction des modèles et l'intégration de connaissances symboliques issues des données ou des utilisateurs. Les contributions de cette thèse montrent qu'il est possible de concilier performance prédictive, explicabilité formelle et interaction humaine au sein d'un même cadre hybride. Elles ouvrent ainsi la voie à des systèmes d'intelligence artificielle plus fiables, plus transparents et mieux adaptés aux besoins des utilisateurs.