



Avis de Soutenance

Madame Nguyen-Duong-Chi TRAN

Informatique et applications

Soutiendra publiquement ses travaux de thèse intitulés

Explications probabilistes pour les prédictions indépendantes du modèle

dirigés par Monsieur Frédéric KORICHE et Monsieur Stefan MENGEL

Soutenance prévue le **mardi 04 novembre 2025** à 14h00

Lieu : UFR des Sciences Jean Perrin Rue Jean Souvraz SP 18 F-62307 Lens Cedex France

Salle : des Thèses

Composition du jury proposé

M. Frédéric KORICHE	Université d'Artois	Directeur de thèse
Mme Céline ROUVEIROL	Université Paris 13	Rapporteuse
M. Frédéric SAUBION	Université d'Angers	Rapporteur
M. Arnaud DURAND	Université Paris Cité	Examineur
M. Stefan MENGEL	Université d'Artois	Co-directeur de thèse
M. Jean-Marie LAGNIEZ	Université d'Artois	Examineur

Résumé :

Expliquer les prédictions réalisées par les modèles d'apprentissage constitue depuis longtemps un défi majeur dans le domaine de l'intelligence artificielle explicable. Récemment, l'explicabilité formelle a émergé comme un sous-domaine essentiel, visant à fournir des explications assorties de garanties mathématiques. Parmi les différentes formes d'explications proposées dans ce cadre, les explications probabilistes se distinguent par leur capacité à concilier concision et précision. Toutefois, le calcul de ces explications probabilistes implique des complexités calculatoires qui dépassent les capacités des solveurs actuels. Pour relever ce défi, nous présentons un cadre théorique unifié permettant de calculer des explications probabilistes avec des garanties d'approximation. Ce cadre combine des éléments d'explicabilité formelle, de théorie de l'apprentissage computationnel et de programmation par contraintes. Pour les modèles de classification, nous reformulons le problème de génération d'explications probabilistes comme une tâche d'apprentissage de termes propositionnels monotones. La minimisation du risque empirique associée peut être résolue via la programmation par contraintes. Dans le cas des modèles de régression, la génération d'explications linéaires probabilistes est formulée comme un problème d'apprentissage de fonctions linéaires parcimonieuses sous contrainte d'hyperplan. Cette tâche de minimisation peut être traitée par programmation en nombres entiers mixtes, et pour les distributions uniformes, elle peut être approximée en temps polynomial grâce à un algorithme de seuillage itératif. Au-delà de ces garanties théoriques, nos résultats empiriques montrent que nos méthodes formelles surpassent les approches heuristiques de pointe telles que Anchors, LIME et MAPLE, en termes de fidélité et d'interprétabilité.