

Avis de Soutenance

Monsieur Aymeric BEAUCHAMP

Informatique

Soutiendra publiquement ses travaux de thèse intitulés

Clustering sous contraintes : intégration incrémentale et active des retours utilisateur

dirigés par Madame Thi Bich Hanh DAO et Monsieur Samir LOUDNI

Ecole doctorale : Mathématiques, Informatique, Physique Théorique et Ingénierie des Systèmes - MIPTIS

Unité de recherche : LIFO - Laboratoire d'Informatique Fondamentale d'Orléans

Soutenance prévue le **lundi 29 septembre 2025** à 10h00

Lieu : LIFO - Bâtiment IIIA, Rue Léonard de Vinci, 45067 Orléans

Salle : des thèses

Composition du jury proposé

Mme Thi-Bich-Hanh DAO	Université d'Orléans	Directrice de thèse
M. Samir LOUDNI	IMT Atlantique	Co-directeur de thèse
Mme Christel VRAIN	Université d'Orléans	Co-encadrante de thèse
M. Pierre GANCARSKI	Université de Strasbourg	Examineur
M. Bruno CREMILLEUX	Université de Caen-Normandie	Examineur
M. Nadjib LAZAAR	Université de Paris-Saclay	Rapporteur
M. Saïd JABBOUR	Université d'Artois	Rapporteur

Mots-clés : apprentissage non supervisé, clustering sous contraintes incrémentales, programmation par contraintes, clustering actif, ensemble de correction minimal, retours utilisateur

Résumé :

Le clustering sous contraintes vise à exploiter les connaissances des experts lors de la tâche de clustering. Les algorithmes traditionnels supposent que toutes les contraintes pertinentes peuvent être fournies dès le départ, ce qui n'est pas réaliste. Dans cette thèse, nous proposons un cadre de clustering sous contraintes incrémentales, où l'utilisateur réagit aux résultats en ajoutant des contraintes progressivement jusqu'à obtenir satisfaction. Nous commençons par introduire un modèle en programmation par contraintes pour réaliser une modification minimale de partition pondérée par la distance des points modifiés à leur nouvelle affectation. Ce modèle assure la continuité des résultats tout au long du processus itératif, un aspect peu étudié dans le clustering interactif. Les résultats expérimentaux sur des jeux de données classiques montrent que, par rapport à l'état de l'art, notre approche exploite plus efficacement les contraintes pour se rapprocher d'une partition cible, tout en préservant la similarité entre les itérations. Nous proposons ensuite un algorithme de clustering actif adapté au cadre incrémental, fondé sur un modèle supervisé appris à partir des retours utilisateur pour suggérer de nouvelles contraintes. Notre algorithme affiche des performances comparables à l'état de l'art tout en étant généralement plus rapide et en apprenant progressivement la partition désirée. Enfin, nous explorons la gestion des conflits entre contraintes pouvant perturber le bon fonctionnement du système. Nous proposons une mesure d'adéquation de contrainte à un modèle d'apprentissage supervisé entraîné sur les retours utilisateur. Il est alors possible de déterminer les contraintes à relâcher en priorité en fonction de leur adéquation.