COLLEGE ED MATHEMATIQUES
DOCTORAL STIC
PAYS DE LA LOIRE

Nantes Université

THÈSE DE DOCTORAT DE

NANTES UNIVERSITÉ

ÉCOLE DOCTORALE N° 641
Mathématiques et Sciences et Technologies de l'Information et de la Communication
Spécialité : *Informatique*

Par
Hugo PEUZET

Classification multi-labels continue avec contraintes de ressources

Thèse présentée et soutenue à Polytech Nantes, le 22 septembre 2026
Unité de recherche : Laboratoire des Sciences du Numérique de Nantes (LS2N – équipe Duke)

Rapporteurs avant soutenance :

Lydia BOUDJELOUD-ASSALA — Professeure des universités, Université de Lorraine - Metz
Jesse READ — Professeur des universités, Ecole Polytechnique

Composition du Jury :

Président :	Prénom NOM	Fonction et établissement d'exercice *(à préciser après la soutenance)*
Examinateurs :	Benoît PARREIN	Professeur des universités, Université de Rennes - ESIR
	Hoel LE CAPITAINE	MCF HDR, Nantes Université - Polytech
Dir. de thèse :	Pascale KUNTZ	Professeure des universités, Nantes Université - Polytech

Invité(s) :

Frank MEYER — Docteur Ingénieur R&D, Orange Innovation Lannion

COLLEGE ED MATHEMATIQUES
DOCTORAL STIC
PAYS DE LA LOIRE

Nantes Université

Titre : Classification multi-labels continue avec contraintes de ressources

Mot clés : classification multi-labels, flux de données, apprentissage continu, oubli catastrophique, apprentissage frugal

Résumé : La classification multi-labels continue en flux consiste à classer "en continu" les instances d'un flux de données décrites par un grand nombre d'attributs et associées à un ensemble de descripteurs. Dans le contexte décisionnel professionnel, l'évolution des techniques d'apprentissage sur des flux de données continus se heurte à plusieurs défis : le besoin d'intelligibilité des processus de calculs, la réduction des coûts des ressources mobilisées, et la minimisation des dérives de concepts qui peuvent engendrer des dégradations de performances et dont la détection et le traitement sont complexifiés dans le cadre multi-labels par les combinaisons de labels. Une nouvelle formalisation de ce problème dans le cadre des données tabulaires a servi de base à la simulation de scénarii de flux de données multi-labels non-stationnaires avec des tâches, permettant d'évaluer les performances prédictives des algorithmes, leur efficacité de transferts entre tâches et leur degré de frugalité. Ce simulateur a été mis en œuvre dans la comparaison d'algorithmes de l'état de l'art à des stratégies de complexité plus faible qui intègrent pour certaines du rejeu de données, mécanisme prometteur pour lutter contre l'oubli catastrophique. En sus de ses apports méthodologiques, ces travaux ouvrent sur des perspectives applicatives en cybersécurité et en intelligence ambiante.

Title: Multi-Label Continual Stream Classification with Resource Constraints

Keywords: multi-label classification, data stream, continual learning, catastrophic forgetting, frugal learning

Abstract: Multi-label continual stream classification consists of "continuously" classifying data stream instances described by a large number of features and associated with a set of descriptors. In the professional decision-making context, the evolution of learning techniques on continuous data streams faces several challenges: the need for intelligibility in computational processes, the reduction of mobilized resource costs, and the minimization of concept drifts, which can lead to performance degradation and whose detection and treatment are complexified in the multi-label framework by label combinations. A new formalization of this problem within the framework of tabular data served as the basis for simulating non-stationary multi-label data stream scenarios with tasks, making it possible to evaluate the predictive performance of algorithms, their task-transfer efficiency, and their degree of frugality. This simulator was implemented to compare state-of-the-art algorithms against lower-complexity strategies, some of which incorporate data replay—a promising mechanism to reduce catastrophic forgetting. Beyond its methodological contributions, this work opens up application prospects in cybersecurity and ambient intelligence.