

Vous faites partie d'EDDA, projet sélectionné et financé par l'ANR dans le cadre de l'Appel à projets trilatéral France Allemagne Japon en Intelligence Artificielle, pouvez-vous nous informer sur les recherches en cours ?

Le projet EDDA (EnhanceD Data stream Analysis) a démarré il y a tout juste quelques mois. Côté français, il rassemble des chercheurs de Toulouse, Paris et Nancy, Greifswald côté allemand, ainsi que des collègues de l'agence japonaise d'océanographie, le JAMSTEC.

Notre objectif est de nous intéresser à l'analyse de flux de données environnementales, tel que le profil océanographique, ou les données climatologiques. Il s'agit d'être capable de comprendre mais aussi de prévoir et d'anticiper les évolutions simultanées de plusieurs phénomènes temporels, qui peuvent être liés de manière complexe les uns aux autres.

Pour aborder cette problématique, nous entendons nous appuyer sur une approche introduite très récemment, la méthode de la signature. Cette méthode, basée sur la théorie des chemins rugueux, permet d'encoder à chaque instant en utilisant ce que nous appelons la signature les dépendances non linéaires entre les différentes quantités temporelles d'intérêt.

Des outils efficaces de calcul de la signature ont été développés par différentes équipes internationales (universités d'Oxford et de Berlin entre autres). De nombreuses applications (parmi lesquelles l'analyse de séries financières, l'astronomie, la détection de pathologies mentales) ont été explorées avec succès car la communauté française et internationale est très active sur le sujet, comme en témoigne la création du Groupe de recherche sur les TRAjectoires ruGueuses (GdR TRAG), soutenu par le CNRS.

Néanmoins, de nombreuses questions liées à l'interprétabilité de la méthode et à l'établissement de garanties statistiques permettant d'évaluer ses performances restent en suspens. Elles constituent un véritable défi auquel nous souhaitons nous attaquer dans le cadre d'EDDA.

Quels sont les thèmes que vous avez envie d'explorer à l'interface entre probabilités, statistiques et apprentissage ?

Les interfaces possibles sont vastes : théorie de l'apprentissage, optimisation stochastique, approches bayésiennes impossible de les lister en quelques mots tant les travaux sont foisonnants dans ce domaine. Je ne vais donc pas m'y risquer !

Actuellement, j'aimerais approfondir deux questions. La première est celle de la prise en compte de l'incertitude en machine learning, qui peut se situer dans la quantité que l'on cherche à prédire ou du bruit présent dans les données collectées. La deuxième question est celle de l'interprétabilité en machine learning, des travaux émergent dans ce domaine. J'aimerais aborder cette question via les applications que je développe avec mes collègues lorrains en linguistique ou en robotique, ainsi qu'avec mes partenaires industriels.



Marianne Clausel

Marianne Clausel est professeure de Statistique à l'Institut Élie Cartan de Lorraine, laboratoire de mathématiques de l'université de Lorraine. Elle a soutenu sa thèse en 2008 en Mathématiques pour le signal à l'université de Paris-Est Créteil (analyse harmonique et ondelettes). En 2011, elle a été nommée au Laboratoire Jean Kuntzmann de l'université Grenoble-Alpes, ce qui lui a permis de développer de fructueuses collaborations en apprentissage statistique avec le Laboratoire d'informatique de Grenoble dont elle est toujours membre associée.

En poste depuis 2017 à l'université de Lorraine, elle travaille sur l'interface entre statistique appliquée, apprentissage et traitement du signal, avec notamment des applications en Traitement automatique des langues dans le cadre du projet de site OLKi.

Elle participe aux activités du Labex AMIES (Agence pour les mathématiques en interaction avec l'entreprise et la société), qui vise à initier et promouvoir les interactions entre académiques et industriels.

Depuis décembre 2020, elle est responsable côté français du projet international EDDA, qui a pour objectif de proposer des méthodes novatrices d'analyse de flux de données via la méthode de la signature.