Stage de M2 pour l'année universitaire 2019-2020

Automatisation d'une procédure de sélection de modèles dans le cadre du Couplage Modèle-Données

Emmanuel Frénod¹
Hélène Flourent^{1,2}

1 Sujet

Le couplage Modèle-Données est aujourd'hui une des pistes les plus prometteuses pour développer des outils d'Intelligence Artificielle dans des domaines où les phénomènes étudiés sont complexes et où des connaissances métier peuvent être utilisées en complément des données collectées. Ces caractéristiques sont notamment rencontrées dans le domaine du Smart Farming.

Cette approche de couplage Modèle-Données, encore peu explorée, se situe à mi-chemin entre l'approche "pure data", très explorée par les spécialistes du Machine Learning et celle du "full model", étudiée en particulier par les EDPistes.

La définition du niveau d'abstraction du modèle est une étape décisive de cette approche. En effet, un niveau d'abstraction important peut permettre d'avoir un outil très léger et facile à implémenter. Néanmoins cette simplicité peut être au détriment de la précision du modèle. À l'inverse un modèle présentant un niveau de détails important, peut contenir un grand nombre de paramètres à déterminer, nécessiter un temps d'apprentissage important et aboutir à des problèmes de surapprentissage. L'objectif est donc de déterminer par des méthodes de sélection, le niveau d'abstraction le plus pertinent en terme de précision du modèle, de temps d'apprentissage et de nombre de paramètres à déterminer.

L'objectif de ce stage est d'automatiser une procédure de sélection de modèles dans le cadre du Couplage Modèle-Données. Au cours d'un projet précédent, un outil a été construit afin d'ajuster un modèle sur des données et de collecter certaines informations associées à celui-ci. L'objectif est donc d'éprouver cet outil, d'augmenter sa robustesse et son niveau de généralité et enfin d'automatiser la procédure complète de sélection de modèles. Ces différentes étapes nécessiteront la manipulation de modèles issus de différentes approches (Couplage Modèle-Donnée, Machine Learning, Deep Learning, *etc.*) et la construction de nouvelles métriques permettant de juger de la pertinence de ces modèles et de les comparer entre eux.

1. Université Bretagne Sud, Laboratoire de Mathématique Atlantique, UMR CNRS 6205, Campus de Tohannic, Vannes, France

2. Neovia, Route de Talhouët, Saint-Nolff, France

2 Profil

- BAC +5 Statistiques / Informatique
- Connaissances : R ou Python, Méthodes de sélection de modèles
- Capacités : Autonomie, rigueur.

3 Informations

- **Durée** : 6 mois à partir de janvier 2020.
- **Localisation** : Université Bretagne Sud, Laboratoire de Mathématique Atlantique, UMR CNRS 6205, Campus de Tohannic, Vannes, France
- **Encadrement** : Emmanuel Frénod, professeur des universités en mathématiques appliquées et Hélène Flourent, chargée de recherche dans une entreprise agroalimentaire.
- **Envoi des candidatures** : CV et lettre de motivation à envoyer par email à emmanuel.frenod@univ-ubs.fr et helene.flourent@univ-ubs.fr