Post-doctoral position 2022 - 2024

Interactive multi-paradigm collaborative learning for time series analysis

CONTEXT

The generation of a quasi-continuous flow of an increasingly massive quantity of data makes it impossible in many domains to perform the labelling phase of supervised learning. Indeed, this task can no longer be carried out by experts as it is too tedious and time-consuming. Moreover, it assumes that the experts already have an a priori definition of the classes (nomenclature, ontology, etc.) that may interest them. However, this is not always available or is only partial. The ANR HERELLES project aims to meet this need by proposing an innovative method of interactive multi-paradigm collaborative learning, which combines supervised and non-supervised methods while allowing interaction with the expert.

This project is based on two complementary aspects, the use of information exchanges between methods, inspired by existing concepts, boosting [3], co-learning [4] for supervised and collaborative [2] for unsupervised, as well as the addition of constraints in this process [1].

OBJECTIVE

The candidate will propose and define novel mechanisms that allow supervised and unsupervised methods to collaborate efficiently to reach a classification consensus. The modalities of information exchange between them will have to be specified. He/she will also have to define a protocol for interaction between the user and the learning methods through the use of constraints. Finally, he/she will have to concretely implement the proposed approaches to allow their testing and validation.

COLLABORATION AND SUPERVISION

The person recruited will be co-directed by Antoine Cornuéjols (AgroParisTech - 50%), a specialist in supervised collaborative learning, and Pierre Gançarski (ICube - 50%), a specialist in collaborative clustering. He/She will actively collaborate with the SDC team of ICube in Strasbourg and more particularly with Antoine Saget, PhD student (1st year) working on the same subject, as well as Baptiste Lafabregue, co-author, with P. Gançarski, of several articles on clustering under time series constraints [1]. He/She may therefore be required to travel there frequently (at the laboratory's expense).

GENERAL INFORMATION

Location: Saclay (AgroParisTech campus, 22 place de l'Agronomie, 91120 Palaiseau)
Duration: One year (renewable once) – starting as soon as possible
Salary: between 2500€ and 2700€ before taxes (brut) monthly according to past experience
Contact: Antoine Cornuéjols antoine.cornuejols@agroparistech.fr and Pierre Gançarski, pierre.gancarski@unistra.fr

EXPECTED PROFILE

- PhD in Computer Science and specialized in Machine Learning/Data Mining.
- Strong knowledge in Data Science and more particularly in standard classification and clustering methods. Experience in using collaborative/ensemble models or constraint integration would be a plus.
- Good verbal (English or French) and written (English) communication skills.
- Interpersonal skills and the ability to work individually or as part of a project team.

TO APPLY

Send an email to antoine.cornuejols@agroparistech.fr included curriculum vitae, list of publications, letter of motivation and contact details of three references. Applications will be accepted until the position is filled.

BIBLIOGRAPHY

[1] T. Lampert, T-B-H. Dao, B. Lafabregue, N. Serrette, G. Forestier, B. Crémilleux, C. Vrain, P. Gançarski, Constrained distance based clustering for time-series: A comparative and experimental study. Data Mining Knowledge Discovery, 32:1663–1707 (2018)

[2] Antoine Cornuéjols, Cédric Wemmert, Pierre Gançarski, and Younès Bennani. Collaborative clustering: Why, when, what and how. Information Fusion, 39 :81–95, 2018.

[3] R.E. Schapire et al. Boosting: Foundations and Algorithms. MIT Press (2012).

[4] Avrim Blum and Tom Mitchell. Combining labeled and unlabeled data with co-training. In Proceedings of the eleventh annual conference on Computational learning theory, pages 92–100, 1998.