MIAI Post-Doctoral position (12 months)

SECTOR**:** Higher Education Institution

LOCATION: France, Grenoble

RESEARCHER PROFILE:

□ *First stage researcher,*

INSTITUTION: Univ. Grenoble Alpes, University of Innovation

One of the major research-intensive French universities, Univ. Grenoble Alpes[1] enjoys an international reputation in many scientific fields, as confirmed by international rankings. It benefits from the implementation of major European instruments (ESRF, ILL, EMBL, IRAM, EMFL*). The dynamic ecosystem, grounded on a close interaction between research, education and companies, has earned Grenoble to be ranked as the 5th most innovative city in the world. Surrounded by mountains, the campus benefits from a natural environment and a high quality of life and work environment. With 7000 foreign students and the annual visit of more than 8000 researchers from all over the world, Univ. Grenoble Alps is an internationally engaged university.

A personalized Welcome Center for international students, PhDs and researchers facilitates your arrival and installation.

In 2016, Univ. Grenoble Alpes was labeled «Initiative of Excellence ". This label aims at the emergence of around ten French world class research universities. By joining Univ. Grenoble Alpes, you have the opportunity to conduct world-class research, and to contribute to the social and economic challenges of the 21st century ("sustainable planet and society", "health, well-being and technology", "understanding and supporting innovation: culture, technology, organizations" "Digital technology").

* ESRF (European Synchrotron Radiation Facility), ILL (Institut Laue-Langevin), IRAM (International Institute for Radio Astronomy), EMBL (European Molecular Biology Laboratory), EMFL (European Magnetic Field Laboratory)

Key figures:

- + 50,000 students including 7,000 international students
- 3,700 PhD students, 45% international
- 5,500 faculty members
- 180 different nationalities
- 1st city in France where it feels good to study and 5th city where it feels good to work
- ISSO: International Students & Scholars Office affiliated to EURAXESS

[1] https://edu.univ-grenoble-alpes.fr/en/

MANDATORY REFERENCES:

PROJECT TITLE: *MIAI @ Grenoble Alpes*
SUBJECT TITLE: Dimensioning AI based algorithms : Bounds, information criterion and applications to times series detection problems.

RESEARCH FIELD (cf mots clefs sur Euraxess Jobs): Information sciences, Geosciences
SCIENTIFIC DEPARTMENT (LABORATORY'S NAME): GIPSA-Lab, ISTERRE
SUPERVISOR'S NAME*: Olivier MICHEL, Michel CAMPILLO*

SUBJECT DESCRIPTION:

The proposed research track is about model assessment and sparse modeling or model complexity in IA representation of time series. Although these issues are commonly addressed via cross validation procedures, they remain extremely costly in general. Furthermore, defining training, validation and tests sets assumes to have sufficient data (or high enough SNR) to split the data into these three parts. Validation steps may be replaced by analytical approaches (such as AIC, MDL, BIC...) or resampling methods, such as bootstrap. It is often claimed that these former criteria have poor performances for data rich problems tackled by AI based algorithms (such as CNN, DNN), for which the number of model parameters may be very high (up to millions). Hastie, Tibshirani and Friedman provide many empirical comparisons in their celebrated book, and show that -at least on the simulations presented in their book- AIC, BIC and SRM seemed to have interesting performances. In a study published later, Cherkassky [1] claimed that VC dimensions and BIC shows similar performances and outperform AIC, in the case of linear and k-NN based regression problems. In the same years, results about VC dimensions for RNNs were derived in the pioneering work of Koiran and Sonntag [2], and is still an active field of research (see e.g. [3]). Alternately another family of approaches were derived, with the purpose of deriving bounds on the generalization error. Probably Approximately Correct (PAC) learners introduced in the early 80's, as well as Rademacher complexity and allow the derivation of these generalization bounds, expressing relations between performances, training set size and model complexity.

We propose to revisit the aforementioned paper in a signal processing perspective, that is to apply these results in the framework of real data from seismic sensors. The purpose of the processing is to detect departure (i.e. anomalies) from a behavior modeled with different approaches, ranging from « classical » multivariate auto-regressive models to more recent Scattering network and Deep recursive neural networks. The dimension of these models is thus expressed in the choice of some key parameters (such as the AR order or the dimension of the cell-state vector in RNNs). 'Classical' VARMA based approaches were recently implemented and their performance in anomaly detection upgraded by the introduction of new statistical tests accounting for spatial and temporal correlations in the sample [4]. Their performances can be analytically studied and their relation to model complexity can be derived. Consequently these methods can serve as benchmark for assessing the performances of non-linear approaches such as RNN, LSTM or GRU, for which clear relation between model complexity, network architecture and performances are not clearly established. The performances and ability of analytical criterion to determine the model complexity will be evaluated and compared, when applied to typical seismic signals.

The candidate should be able to have or develop knowledge in AI for times series learning and processing. He will conduct original scientific research by designing novel approaches and understanding of AI based algorithm from both theoretical and applied perspectives. An interest for geophysical problems is encouraged. Programming skills in Matlab or Python will be appreciated.

[1] Cherkasski V, Ma Y. : Comparison of model selection for regression, in Neural Comput. 2003 Jul; 15(7):1691-1714
[2] Koiran P; and Sonntag E.D. : Vapnik Chervonenkis dimension of recurrent neural networks, Discrete and Applied Math 86, 1998 : 63-79
[3] Scarselli F., Tsi A.C., Hagenbuchner M. : The VC dimension of graph and RNNs, Neural Networks 2018, 108:248-259.
[4] Elbouch et al. : A normality test for Multivariate Dependant Samples, ArXiv:2109.08427 **[stat.ME]**

AIC = Akaike Information Criterion
BIC = Bayes Information Criterion
CNN = Convolutional Neural Networks
DNN = Deep neural Networks
GRU = Gated Recurent Units
LSTM = Long Short Time Memory
MDL = Minimal Description length
RNN = recurent Neural Networks
SRM = Structural Risk Minimization
VARMA = Vectorial Auto Regressive and Moving Average
VC dim = Vapnik Chervonenkis dimension

ELIGIBILITY CRITERIA
Applicants must hold a PhD degree at latest on the starting date of the post Doc (Jan 1st or Feb 1st, 2022)

Applicants will have to send an application letter in English and attach:
- Their last diploma
- Their CV
- A short presentation of their scientific project (2 pages max)
- Letters of recommendation are welcome.

Address to send their application: olivier.michel@grenoble-inp.fr, michel.campillo@univ-grenoble-alpes.fr