**Postdoc (24 months)**
**Host organization** :Unité d'Informatique et d'Ingénierie des Systèmes (U2IS), ENSTA Paris
828, Boulevard des Maréchaux, 91762 Palaiseau
**Project members** : Gianni Franchi (U2IS), Nicolas Bousquet (EDF),
gianni.franchi@ensta-paris.fr, nicolas.bousquet@edf.fr

# Postdoc in anomaly detection in noisy multivariate time series

The U2iS laboratory of ENSTA Paris at Institut Polytechnique de Paris is looking for a motivated and enthusiastic young researcher to work on the detection and prediction of anomalous behaviors in multivariate time series involving noisy observations of environmental variables and industrial sensors. Founded in 1741 ENSTA Paris is the oldest "Grande Ecole" in France and is located in Palaiseau in the south of Paris.

We seek a research **post-doc** interested in **anomaly detection in time series**, characterized by strong asymetry and noise. Anomaly detection is an important topic that can be done in a supervised, semi-supervised, or unsupervised way. The future postdoc should have a statistical and machine learning background.

The candidate should have knowledge of machine learning techniques. Computer science/ or mathematician profiles are also welcome to apply. The future applicant will be given the opportunity to apply to excellent international publications (ICML, NIPS, ICLR,CVPR, ICCV, ECCV,TPAMI, IJCV).

**Goal :** Anomaly detection in time series is an essential field. Time series are lists of observations at different times linked by the time. Hence the time can provide us with a piece of crucial information to detect anomalies. In [1] Hawkins *et al.* defined anomalous data as : "an observation that deviates so much from other observations that it raises the suspicion that a different mechanism generated it". Detecting these anomalies is crucial for multiple industrial tasks. Using deep learning could be interesting, yet that is not straightforward. Indeed we often have a lot of in-distribution data but very little out-of-distribution data. This can distort neural networks' training that tends to overfit the in-distribution. One first solution consists in using regularization and better working on the loss of the neural network. This can be interesting, yet, all these approaches can still perform overfitting when we see new anomalies. An approach quite used in the industrial context [2, 3, 5, 4] consists in learning an interesting representation of the neural network. Once this representation has been learned, classical anomaly detection algorithms are applied to it. Fundamentally, it appears relevant to come to the foundations of learning and updating, through Bayes' theorem, and to accompany this work with thoughts on the explainability and limitations of the approach.

This postdoc is in collaboration with EDF, which will provide us with anomaly data. Thus, we already have a dataset, and we want to see how we can detect anomalies in it. We will not limit ourselves to this dataset. However, it is important to note that we expect results on this dataset. Depending on how this work evolves, some other cases that involve a mixture of noisy observations and correct (sensor) observations on production plants might be used.

Depending on the candidate's profile, the future worker might focus more on the mathematical aspects of anomaly detection or of representation learning.

**Your task :**

- You will be in charge of projects and activities related to the research in Anomaly detection.
- You will help in improving an Anomaly detection database.
- You will lead the research work.
- You might supervise internship students.
- You will publish the results of your research in scientific international conferences and journals.

**Desired profile :**

- PhD or Master degree in Machine Learning, Statistics, Computer science or related fields
- Relevant scientific track record on major computer vision conferences/journals (NIPS, ICLR, ICML, CVPR, ICCV, ECCV, TPAMI, IJCV, etc.)
- Experience on machine learning and, in particular in sklearn framework
- Experience on DNN models and, in particular in Pytorch framework
- Good communication skills and ability to cooperate
- Proficient in English language (written and oral)

**To formally apply, please send us an email with your resume and your reference letter.**

[1] Hawkins, Douglas M. Identification of outliers. Vol. 11. London : Chapman and Hall, 1980.

[2] Pang, Guansong, et al. "Deep learning for anomaly detection : A review." ACM Computing Surveys (CSUR) 54.2 (2021) : 1-38.

[3] Defard, Thomas, et al. "Padim : a patch distribution modeling framework for anomaly detection and localization." International Conference on Pattern Recognition. Springer, Cham, 2021.

[4] Liznerski, Philipp, et al. "Explainable deep one-class classification." arXiv preprint arXiv :2007.01760 (2020).

[5] Roth, Karsten, et al. "Towards total recall in industrial anomaly detection." arXiv preprint arXiv :2106.08265 (2021).