



Internship offer for 2024

This internship could lead to a PhD thesis in the [SOUND.AI](#) program

Few-shot Learning: application to electricity consumption load curves forecasting

The project

EDF develop services for both individuals and business customers. In many of these applications time series forecasting is important as the data is electrical load curve (electric consumption measurement at regular time steps).

EDF have developed several forecasting models for applications like:

- forecasting end of the month consumption for industrial, large consumers, residential to have accurate monthly payment (to reduce regularization bills)
- forecasting daily/monthly consumption to make recommendation for tailored tariffs (most relevant service to customers according to their profile and consumption) or relevant action to reduce energy consumption (sobriety)
- forecasting aggregated consumption of EDF customers at national level or for sub-aggregates of customers to ensure balance with production and network stability

EDF aims to extend the actual forecasting models for (new) customers with little or no historical data without loss of accuracy on the forecasting results and then the quality of the service provided based on them. Moreover, the actual methods whether being statistical (GAM, generalized additive models [1]) or deep learning-based can be significantly affected in atypical situations: disruptions, rare events (cold/heat waves, incentives for sobriety via the ecowatt system) for which few or no historical data are available.

These issues can be addressed as few-shot learning [2,3,4] problem and, could be coupled with multi-task learning methods [5,6] which allow to jointly infer consumption evolutions at individual and aggregate levels for more accurate prediction.

This internship will then focus on few-shot learning for time series forecasting to solve the issues above. For that the expected planning is:

- Literature review on few-shot learning methods for time series forecasting.
- Implementation of a selected method on internal EDF load curves data.

If there is enough time, start literature review on multi-task learning.

Notes:

This work is preliminary work for a PhD proposal carried out by EDF, Sorbonne and UCL/Inria in the SOUND.AI program:

<https://soundai.sorbonne-universite.fr/dl/subjects/s/ff30ae/r/U2Z6mr5JToqTkGD-u16TuA>

The candidate is expected to fall into the profile requirements below and will be asked to

apply to the PhD program (more details will be given in interview).

The selected intern will benefit from the informed input of the academic researchers who are associated with the PhD proposal.

Profile requirements:

- Master's student in Data Science, Machine Learning, Statistics, or related fields
- Must not have reside or carried out main activities (work, studies) in France for more than 12 months the past 36 months
- English proficiency test results: TOEIC 800 or TOEFL 79 or CEFR C1 or IELTS 6.5
- Good knowledge of statistical/machine learning for data analysis
- Good programming skills in Python or R
- Curiosity, creativity, collaborative spirit
- Clear and effective verbal and written communication skills
- Knowledge of time series analysis is desired but not mandatory

Practical details on the internship

- The internship will take place at
EDF R&D, Department SEQUOIA
EDF Lab Paris-Saclay, 7 boulevard Gaspard Monge, 91120 Palaiseau, France
- Duration: 4-6 months
- Gratification according to education level
- To apply, please send a CV to ghislain.agoua@edf.fr and amaury.durand@edf.fr

References

[1] Wood, S. N., Goude, Y., & Shaw, S. (2015). Generalized additive models for large data sets. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 64(1), 139-155.

[2] Iwata and Kumagai (2020). Few shot learning for time series forecasting, preprint <https://arxiv.org/pdf/2009.14379.pdf>

[3] Lukas Brinkmeyer, Rafael Rego Drumond, Johannes Burchert, Lars Schmidt-Thieme, Few-Shot Forecasting of Time-Series with Heterogeneous Channels <https://arxiv.org/abs/2204.03456>

[4] Victor Tran and Anand Panangadan. 2022. Few-shot Time-Series Forecasting with Application for Vehicular Traffic Flow. In 2022 IEEE 23rd International Conference on Information Reuse and Integration for Data Science (IRI). IEEE Press, 20–26. <https://doi.org/10.1109/IRI54793.2022.00018>

[5] Leroy, A., Latouche, P., Guedj, B., & Gey, S. (2022). MAGMA: inference and prediction using multi-task Gaussian processes with common mean. *Machine Learning*, 111(5), 1821-1849.

[6] Liu, C. L., Tseng, C. J., Huang, T. H., Yang, J. S., & Huang, K. B. (2022). A Multi-task Learning Model for Building Electrical Load Prediction. *Energy and Buildings*, 112601.