Offre de Post-doctorat
Optimization and forecasting of logistic flows

Overview

Califrais is a startup whose mission is to use AI to decarbonize the food supply chain. The startup deploys its technology at Rungis International Market, with the mission of optimizing food flows of the world's largest fresh produce market. To do this, it has developed its own algorithms, thanks to its partnership with the academic world, in particular with the Laboratoire de Probabilités, Statistique et Modélisation (LPSM) at Sorbonne Université and Université Paris Cité.

Califrais has won the "Logistique 4.0" call organized by the ADEME as part of the "France 2030" plan. The aim of this project is to take the solutions developed by Califrais in Rungis to an industrialization phase across the entire network of French wholesale markets, in order to leverage the benefits and impact of the technologies being developed.

Indeed, the 22 French food wholesale markets are key multimodal hubs for national food security, resilience and sovereignty, as well as for the organization of urban logistics. However, the logistics flows between these markets, as well as their economic and environmental costs, remain largely under-optimized as they are neither mutualized nor centralized.

This position is funded within this project, and has the general objective of developing algorithmic solutions for the optimization of food flows through the network of French wholesale markets. The problems can be divided into 3 general topics :

- development of forecasting models adapted to this application (multivariate, high-dimensional, noisy and sparse signal over a graph)
- development of optimization tools for inventory problems
- development of optimization tools for vehicle routing problems

These 3 topics can interact with each other (for example, take forecasting errors into account in the inventory optimization module).

Missions

As a post-doc in the joint LPSM-Califrais research team, you will be expected to:

- Conduct research on problems of the Logistique 4.0 project
- Participate in the creation of machine learning models and data modeling
- Contribute to the creation of robust, scalable and high-performance code to test and compare models and reproduce results.
- Understand, challenge and, if necessary, implement or modify the scientific research carried out by the research team of Califrais
- Co-supervise internships and collaborate with the PhD students at Califrais

The post-doc will be co-supervised by Claire Boyer (Maître de conférences HDR, Sorbonne Université), Gérard Biau (Professor, Sorbonne Université), and Adeline Fermanian (Head of research, Califrais). It will be based both in Sorbonne Université (campus Jussieu, Paris 75005) and Califrais (Paris 75010).

The specific scientific subject is flexible and can be adapted to the candidate's area of expertise. As an example, we detail below one research direction that could be taken for the forecasting problem, but other options can be explored.

Forecasting food flows on a grid

At Califrais, we have a catalog of products (about 10 000), and a 7-year history of daily orders from that catalog. Given this history, we need to be able to forecast future sales, typically within a horizon of 4 weeks. Such forecasts are critical for inventory management and scaling operations.

This means that we have to forecast 10 000 time series that are not independent and have a hierarchical structure: these 10 000 time series at the product level can be grouped into categories, and categories of categories… For example, a low-level time series, corresponding to a product ID, will be a product of a specific brand and a specific packaging, such as "Piment Jalapenos Vert", a first category will be "Légumes", a finer one will be "Poivrons et piments", and a third one will be "Piment Jalapenos". This frames our problem into the setting of hierarchical time series forecasting.

We are currently able to build good forecasters at different levels of the hierarchy. For example, we can have a forecaster of the total demand, or a forecaster at the lowest level. Now we want to be able to share information across the different levels of the hierarchy. Indeed, different levels contain different types of information: at the highest level, the total demand, it will be easy to have information about trends, general patterns such as holidays, annual seasonality… However, we will not be able to predict effects that are specific to some categories of products. Moreover, for some products, the demand will be stable and regular, while for others it will have a lot of noise. We would like to be able to leverage the information we have learned about products that are easier to predict to products that are harder to predict.

In other words, we want to create a mixture of different forecasts, in order to improve the general accuracy. Several lines of work can be explored, such as neural networks and attention-based methods (Lim et al., 2021 ; Mancuso, 2021 ; Wang et al., 2022) or ensemble methods (Goehry, 2019 ; Leprince, 2023).

All these methods must be adapted to the hierarchical aspect (time series at the parent level are supposed to be the sum of the children) and to the specific aspects of our application (daily data, a lot of noise and stochasticity…).

A second step in the project would be not only to provide forecasts for the 10 000 signals, but also reliable confidence intervals. Indeed, the forecasts are used as inputs to a stock optimization algorithm. If given a level of confidence, then the stock optimization algorithm could have a more conservative approach for the uncertain products, and be allowed to play more on the products for which we are confident.

Profile

A PhD in mathematics or computer science is required, ideally in machine learning or optimization.

If you are interested in this offer, send a CV and motivation letter to adeline.fermanian@califrais.fr

We are committed to diversity and inclusion. We strongly encourage people from under-represented groups to apply.

The position is funded for a period of 2 years, with a start date as soon as possible.

References

Goehry, B., Goude, Y., Massart, P., & Poggi, J. M. (2019). Aggregation of multi-scale experts for bottom-up load forecasting. *IEEE Transactions on Smart Grid*, 11(3), 1895-1904.
Leprince, J., Madsen, H., Møller, J. K., & Zeiler, W. (2023). Hierarchical learning, forecasting coherent spatio-temporal individual and aggregated building loads. *arXiv preprint arXiv:2301.12967*.
Lim, B., Arık, S. Ö., Loeff, N., & Pfister, T. (2021). Temporal fusion transformers for interpretable multi-horizon time series forecasting. *International Journal of Forecasting*, 37(4), 1748-1764.

Mancuso, P., Piccialli, V., & Sudoso, A. M. (2021). A machine learning approach for forecasting hierarchical time series. *Expert Systems with Applications*, 182, 115102.
Wang, S., Zhou, F., Sun, Y., Ma, L., Zhang, J., & Zheng, Y. (2022, November). End-to-End Modeling of Hierarchical Time Series Using Autoregressive Transformer and Conditional Normalizing Flow-based Reconciliation. In *2022 IEEE International Conference on Data Mining Workshops (ICDMW)* (pp. 1087-1094).
Salinas, D., Flunkert, V. Gasthaus, J., Januschowski, T. (2020). DeepAR: Probabilistic forecasting with autoregressive recurrent networks. *International Journal of Forecasting* 36(3), 1181-1191.
Challu, C., Olivares, K. G., Oreshkin, B. N., Ramirez, F. G., Canseco, M. M., & Dubrawski, A. (2023). Nhits: Neural hierarchical interpolation for time series forecasting. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 37, No. 6, pp. 6989-6997).