On possible links between hierarchical forecasting and interpretable machine learning

IIF Workshop on Forecast Reconciliation

Margaux Brégère

September 2023 - Monash University Prato, Italy

LABORATOIRE DE PROBABILITÉS STATISTIQUE & MODÉLISATION

Motivation and notations

Interpretable Machine Learning

How interpreting a deep neural network or random forest? Complex machine learning models achieve high accuracy, but do they make sense? Is it possible to open up black box models and explain how they predict future values?

As more and more businesses (with industrial risk, ethical, fair or ecological concerns) rely on machine learning, interpretability is a key issue.

Forecasting problem:

Building a model $\hat{f} : \mathbb{R}^p \to \mathbb{R}$ to predict the future value of a real random variable Y based on p features $X = (X_1, \ldots, X_p)$

Interpretation with the Shapley values $\hat{\phi}_j$:

Decomposition of the prediction $\hat{f}\left(X_{\text{new}}\right)$ into a sum of feature effects $\sum_{j=1}^{p} \hat{\phi}_j\left(X_{j,\text{new}}\right)$

Shapley values - a short detour through game theory

Feature value $X_j = x_j \Leftrightarrow$ Player j

Forecast $\hat{f}(x) = \hat{f}(x_1, \ldots, x_p) \Leftrightarrow$ Payoff of a collaborative game $v\big(\{1,\ldots,p\}\big)$

How to fairly distribute the payoff of a collaborative game among the players ?

$$\phi_{j,v} = \frac{1}{\text{number of players}} \sum_{\text{coalition including } j} \frac{\text{contribution of } j \text{ to coalition}}{\text{number of coalition including } j} = = \frac{1}{p} \sum_{S \subseteq \{1,2,\ldots,p\} \setminus \{i\}} \frac{1}{\binom{p-1}{|S|}} \big(v(S \cup \{i\}) - v(S)\big)$$

For machine learning interpretation, the contribution of the coalition $x_J = \{x_j \,|\, j \in J\}$ where $J \subseteq \{1,2,\ldots,p\}$ is defined as the prediction for feature values x_J in that are marginalised over features that are not included in set J:

$$v(x_J) = \mathbb{E}_{X_{\bar{J}}}\big[\hat{f}(x_J, X_{\bar{J}})\big] = \int \hat{f}(x_J, X_{\bar{J}}) \mathrm{d}\mathbb{P}_{X_{\bar{J}}}$$

→ approximation with Monte-Carlo sampling to calculate the Shapley values

Hierarchical forecasting for Shapley values

Shapley values calculation is time consuming (only approximate solutions are feasible)

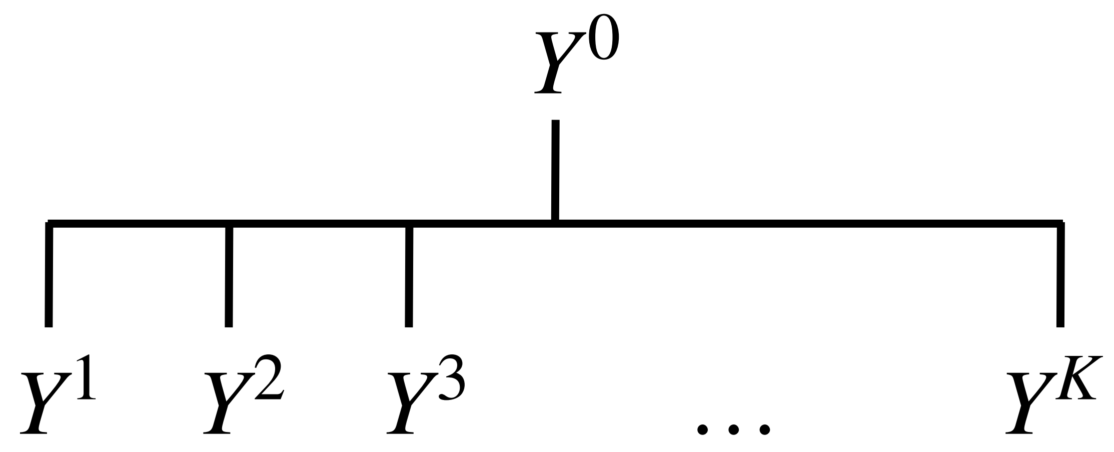
The decomposition $\hat{f}(X_{\text{new}}) = \sum_{j=1}^{p} \hat{\phi}_j(X_{j,\text{new}})$ introduces a hierarchy

Could we use hierarchical forecasting - such as top-down approach - to

- accelerate the calculation?
- provide Shapley value forecasts for new points?

Shapley values for hierarchical forecasting

$$Y_i^0 = \sum_{k=0}^{K} Y_i^k$$



With correct reconciliation:

$$\sum_j \hat{\phi}_j^0(X_{j_0,\text{new}}) = \sum_k \sum_{j_k} \hat{\phi}_{j_k}^k(X_{j_k,\text{new}}^k)$$

But can we say a bit more about it?

- what happens when feature values are the same for every node in the hierarchy: $x_{j,i}^0 = x_{j,i}^k$
- and when the feature (temperature for e.g.) are the same for each node in the hierarchy, but with different values (one per region): $x_{i,j}^0 = g(x_{i,j}^1, \dots, x_{i,j}^K)$
- do the model types (linear regression, random forest, etc.) play an important role?

Some intuitions:

$$\hat{\phi}_j^0(X_{j,\text{new}}) = \sum_k \hat{\phi}_j^k(X_{j,\text{new}})$$

seems a natural result since Shapley values are unique and additive (Game Theory)

As in ML, we only get Shapley value approximations, it is perhaps possible to obtain statistical results for at least some model families (linear, GLM, GAM, neural networks?)

$$\hat{\phi}_j^0(X_{j,\text{new}}) \stackrel{?}{=} \sum_k \hat{\phi}_j^k(X_{j,\text{new}}^k)$$

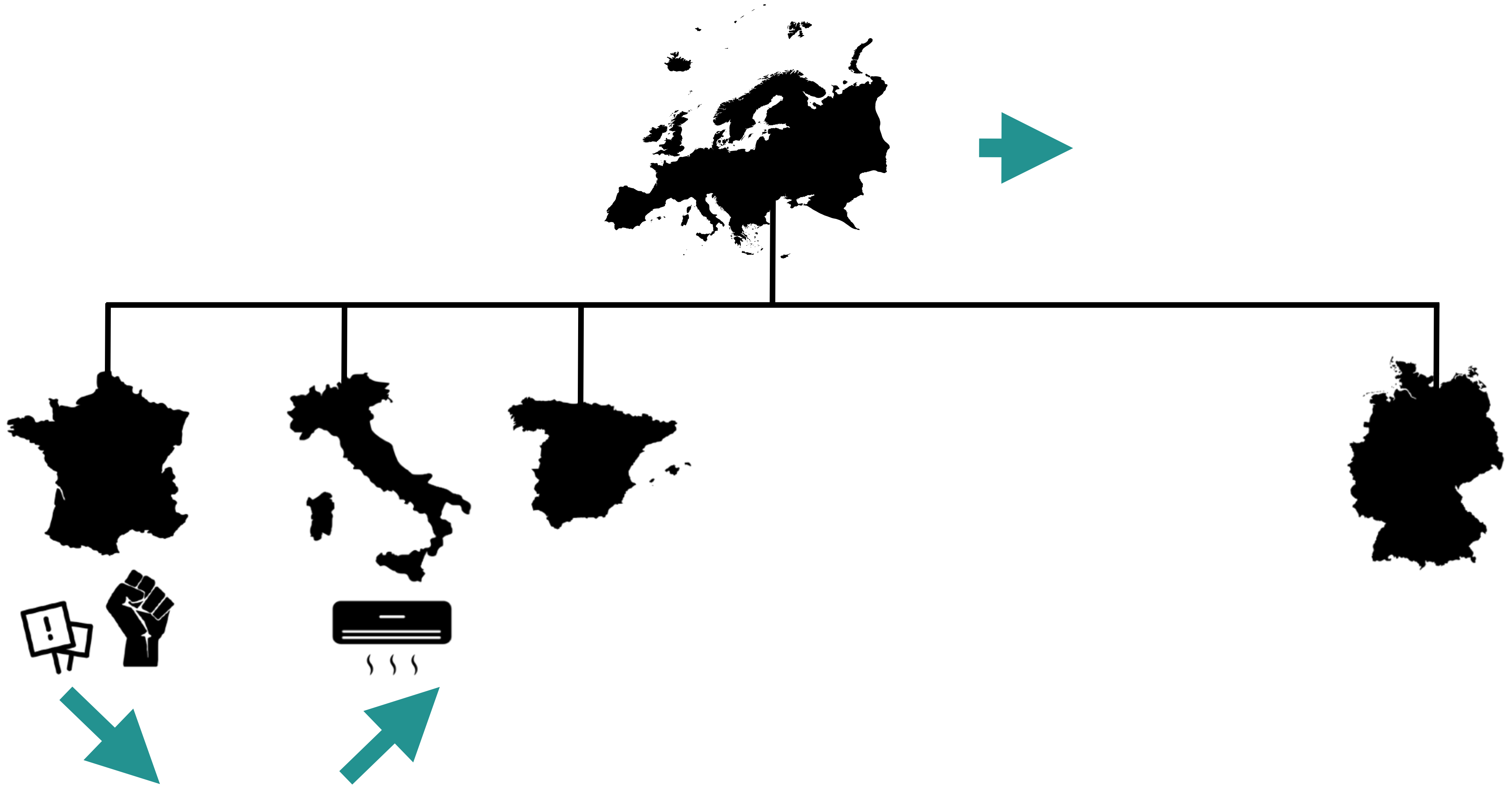
seems a bit more tricky (players are different in the $K+1$ games) but still intuitive

In my opinion, it must strongly depend on the existence of a link between the feature value at the top of the hierarchy and the ones at the bottom: $x_{i,j}^0 = g(x_{i,j}^1, \dots, x_{i,j}^K)$

A more high-level question

What kind of interpretation do we want for global forecasts based on the interpretations of local forecasts?

An example: Forecasting European electricity consumption



« The drop in French consumption caused by the strike
was offset by higher consumption during the Italian heatwave »
or « Business as usual »

Thank you for your attention

Discussion

Comments