

XAI on Time Series Forecasting

Marco Zanotti

University Milano-Bicocca

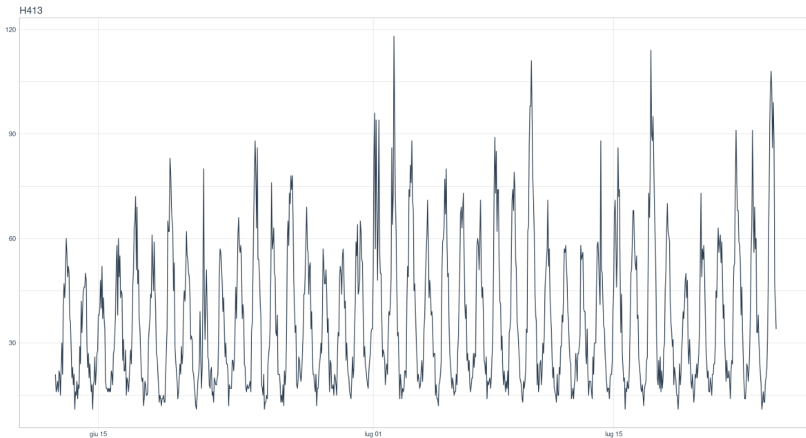


Contents

1. Data
2. Modelling
3. XAI
4. Conclusions

1. Data

M4 Competition



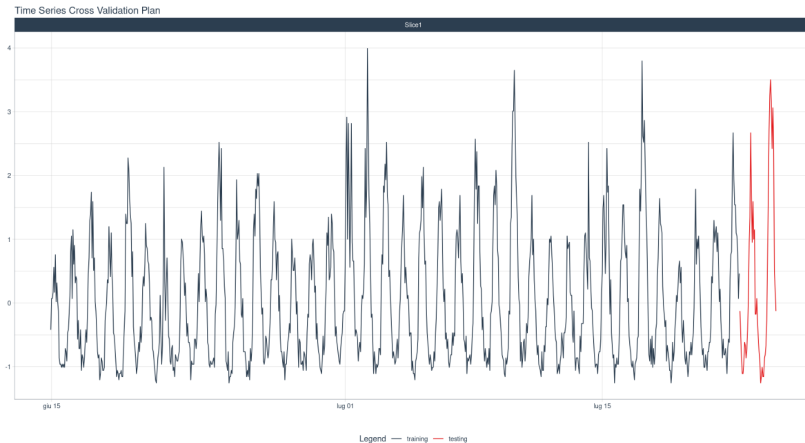
Feature Engineering

In order to be able to use data in a machine learning model, we need to create features that can be used as predictors.

This step is extremely relevant in time series data, since we need to create **features that are able to capture the time dynamics of the data**.

Here, I used lags, rolling features, calendar features and fourier series, obtaining a total of **24 features**.

Train-Test Split



2. Modelling

AutoML with H2O

h2o automatically estimates and tests **6 different ML algorithms**:

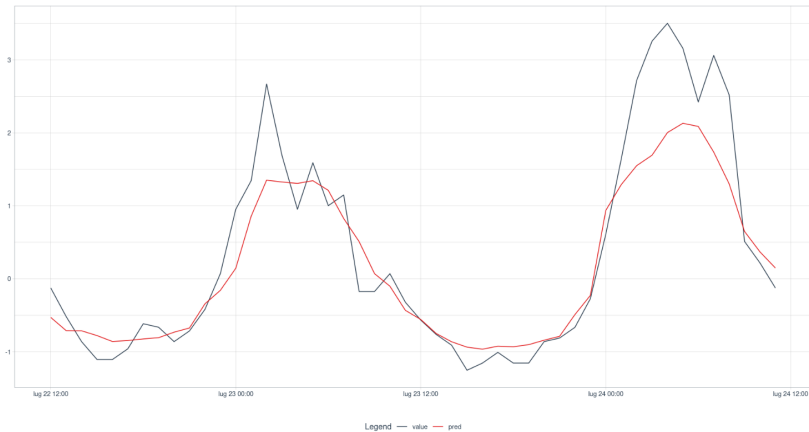
- ▶ DRF (This includes both the Distributed Random Forest (DRF) and Extremely Randomized Trees (XRT) models)
- ▶ GLM (Generalized Linear Model with regularization)
- ▶ XGBoost (XGBoost GBM)
- ▶ GBM (H2O GBM)
- ▶ DeepLearning (Fully-connected multi-layer artificial neural network)
- ▶ StackedEnsemble (Stacked Ensembles, includes an ensemble of all the base models and ensembles using subsets of the base models)

AutoML with H2O

```
model_h2o_automl <- h2o.automl(  
  y = target, x = x_vars,  
  training_frame = train_h2o,  
  max_runtime_secs = 120,  
  max_runtime_secs_per_model = 30,  
  max_models = 50,  
  nfolds = 5,  
  sort_metric = "rmse",  
  verbosity = NULL,  
  seed = 123  
)
```

Best Model

The best model ends up to be a GBM.



3. XAI

XAI in Time Series

In the context of time series forecasting, being able to understand the model's predictions is of paramount importance for 3 main reasons:

1. **Trust:** to be able to trust the model's predictions, especially when it is used to make important future decisions.
2. **Improvement:** to be able to understand the model's weaknesses to improve it.
3. **Combine:** usually model's predictions are combined with human judgement, so understanding what causes such predictions is crucial for business experts to adjust their forecasts.

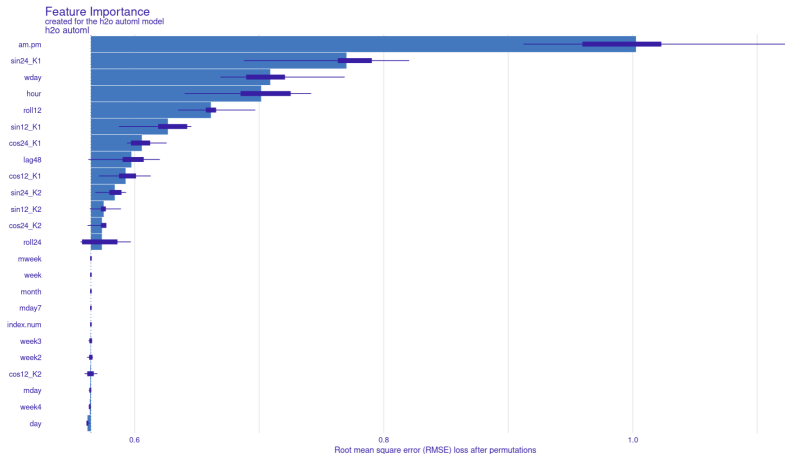
DALEX

To perform a XAI analysis on the automatic black-box model the **DALEX** package, from Dr. Why AI, is used.

DALEX is a XAI framework which allows to easily adopt several model agnostic explainability techniques, such as:

- ▶ Feature Importance
- ▶ Partial Dependence
- ▶ Break Down
- ▶ Shapley Values
- ▶ LIME
- ▶ Ceteris Paribus and
- ▶ Stability Analysis.

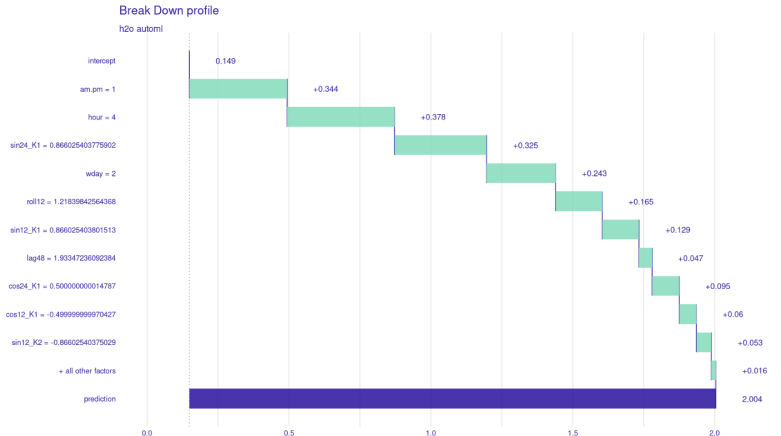
Global - Feature Importance



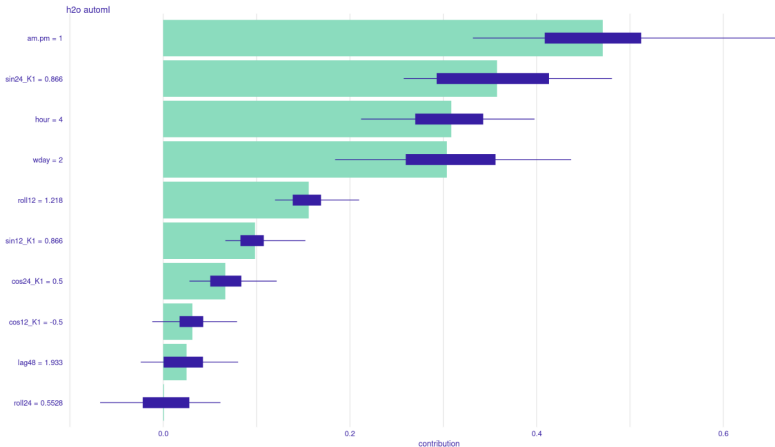
Global - Partial Dependence



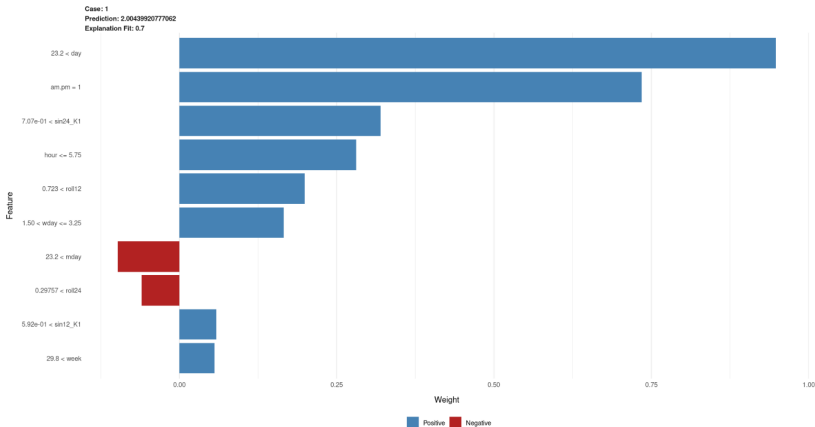
Local - Break Down (observation: 2017-07-24 04:00:00)



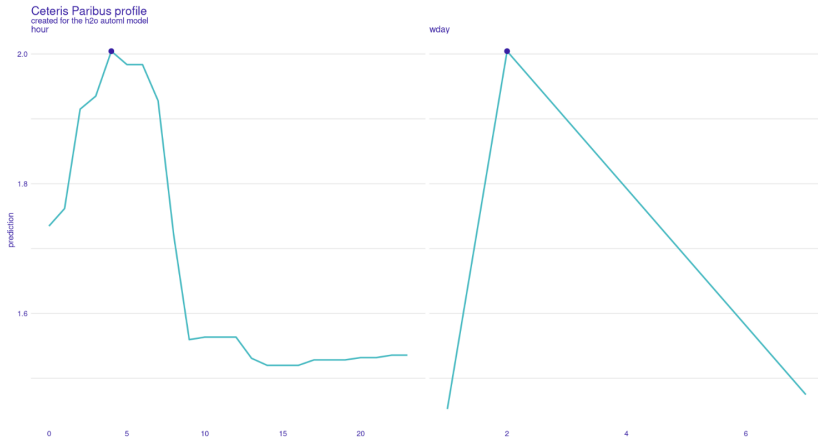
Local - Shapley Values (observation: 2017-07-24 04:00:00)



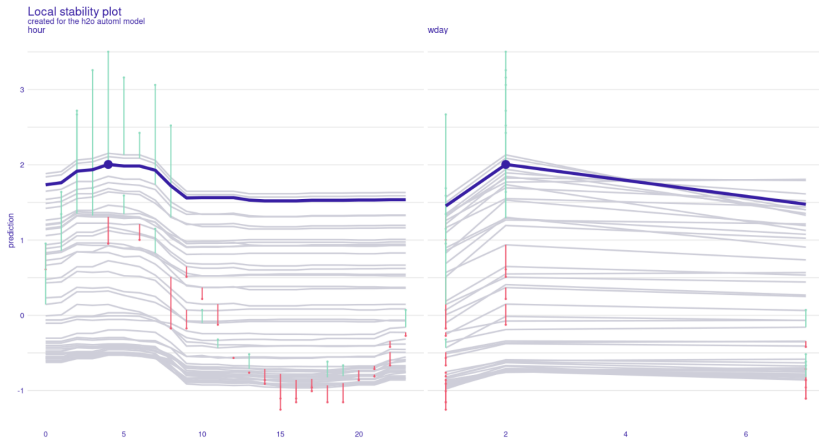
Local - LIME (observation: 2017-07-24 04:00:00)



Local - Ceteris Paribus (observation: 2017-07-24 04:00:00)



Local - Stability Analysis (observation: 2017-07-24 04:00:00)



4. Conclusions

Conclusions

Thank you!