IF YOU WANT TO REBEL
DO IT WITH PURPOSE

Finalist: Maximillian Weil

Abstract

Day-ahead Time Series Forecasting of the Electricity Consumption on the Low Voltage Distribution Grid

An energy revolution is on its way, driven by the realization that our current way of producing and consuming energy is not sustainable. New technologies such as electric heat pumps and electric vehicles bring a solution to this problem. But these changes also bring a challenge to our electricity grid. As these new technologies require a lot of electricity, risks of congestion of the electricity grid at the level of the Low Voltage Distribution Grid (LVDG) become more frequent.

Predicting potential congestion beforehand can allow Distribution System Operators (DSO), such as Fluvius, to take precautionary measures to avoid this. These predictions require both a good knowledge of the electricity grid layout and accurate day-ahead forecasts of electricity consumption at the household level. This thesis aims at improving the forecast side, by implementing Machine Learning (ML) models. The ML models can be used to learn the electricity consumption behaviour of different households from historic consumption data and to derive which external parameters, such as the weather or traffic, influence this behaviour.

The Extreme Gradient Boosting (XGB) regression model is chosen for its high accuracy and interpretability, as proven by the many prediction challenges won by implementations using XGB. The model is implemented as a forecasting algorithm on an openly available data set, provided by Fluvius. The dataset contains the quarter-hourly electricity consumption data for 100 households in Belgium for the year 2016. Unfortunately, the XGB model cannot predict the highly random peaks, present in the consumption data. This is due to the so-called 'double peak penalty', which follows from the fact that a peak predicted with a small time delay, will generate both the error of the wrong peak prediction and the actual peak that wasn't predicted. This leads traditional ML models to predict baseline consumption only. However, for Distribution System Operators, the peak behaviour is key for preventing congestion, as this drives the maximal electrical consumption observed in the LVDG. This motivates the adaption of the model to allow for a small time lag between predicted and actual peak, implemented through a custom objective function on which the model is optimized. This shows promising results for some households. But for others, the benchmark model, based on maximal autocorrelation of historic consumption data outperforms the custom XGB model.

Readily available external variables, that could influence electricity consumption are added to the XGB model input for consumption forecasting. These variables include temporal, weather and traffic data. The game theory based SHAP methodology of Explainable AI is then used to understand which input variables influence the output the most. This is done to extract a general feature importance, that can be used to explain external variables' influence. Additionally, the SHAP values are used to understand the peak behaviour by focusing on peaks predicted by the models. Interestingly, the traffic data showed to be influential for the predicted consumption of three different households. This is expected as people tend to use a lot of electricity before the morning rush hour and after the evening traffic congestions.

This thesis presents tools that Distribution System Operators can use to better understand consumption behaviour at the household level. Additionally, the results of applying these tools on electricity consumption of multiple different households are given in the thesis. This analysis shows very different results for different households, emphasising the variability in household electricity consumption. This allows for a potential classification of household electricity consumption based on different predictability factors.