Case 1. Energy Forecasting

ISDS 7075 Business Forecasting

As electricity cannot currently be stored in large amounts, supply and demand must always be matched or balanced by system operators. That is why accurate energy demand forecasting plays a key role in this.

The liberalization of the electricity industry and introduction of competition in recent decades has introduced wholesale electricity markets where suppliers, generators, and traders buy and sell energy. Electricity retailers, investment banks, and large energy users trade large quantities of electricity and make deals that cover timescales ranging from several years ahead to on-the-day spot trading.

Short term electrical load forecasting is critical in ensuring reliability and operational efficiency for electrical systems. With an influx of monitoring data and the growing technical complexity of the grid, there is a great interest and need for accurate forecasting in electricity planning.

The capstone class project involves a hierarchical forecasting problem of electrical loads. You are required to predict hourly loads (in kW) for a utility company in 15 different geographical regions. Electricity demand is subject to a range of factors such as weather conditions, calendar effect, economic activity, and electricity prices. However, you are required to use only temperatures and calendar information in the competion. (You are not supposed to attempt to get additional *external* information.)

In the project, your task is to develop a time series forecasting model that takes temperature and calendar information as inputs, and predicts electrical loads as the output.

1. Data Set

The dataset contains hourly load measurements for 15 geographic sub-areas (e.g., residential and industrial zones) from January 2015 to June 2019. It also includes hourly temperature readings from 9 weather stations in that region, but we are unaware of the location of the 15 zones and the 9 weather stations. In particular, we do not know which weather stations are located in or near which zones..

Given actual temperature history, the 8 weeks below in the load history are set to be missing and are required to be "backcasted". (It's OK to use the entire history to backcast these 8 weeks.)

- 2016/3/6 ~ 2016/3/12
- 2016/6/20 ~ 2016/6/26
- 2016/9/10 ~ 2016/9/16
- 2016/12/25 ~ 2016/12/31
- 2017/2/13 ~ 2017/2/19

- 2017/5/25 ~ 2017/5/31
- 2017/8/2 ~ 2017/8/8
- 2017/11/22 ~ 2017/11/28.

In addition, you need to "forecast" hourly electrical loads from 2019/6/23 to 2019/6/29. No actual temperatures are given for this week.

2. Predictive Modeling

"Forecasting" is the process of exploring the future events that have not been observed or determined. On the other hand, "backcasting" typically refers to the process of exploring the past events given the information known to date.

In the project, you need to "backcast" the electrical loads for each of the eight missing weeks in the past. You may use other load measurements the hourly temperatures. For each problem, you also need to aggregate their prediction to estimate the total load over all 15 different regions.

Assuming that today is June 22, 2019, you also need to "forecast" the energy consumption for the next one week. In such a case, the forecasts are obviously based on the future temperatures that should be also predicted.

There are many types of time series forecasting models, such as multiple linear regression, exponential smoothing, ARIMA models, dynamic regression, and combined forecasts. You may choose any forecasting models with many bells and whistles.

3. Performance Evaluation

The forecasting accuracy used in the competition is a weighted root mean squared error, where the weights are assigned as follows:

- Each hour of the 8 backcasted weeks at regional level: 1
- Each hour of the 8 backcasted weeks at system level: 15
- Each hour of the 1 forecasted week at regional level: 8
- Each hour of the 1 forecasted week at system level: 160

In addition to the forecasting accuracy, your report will be evaluated based on the clarity of documentation, rigors of the approach, interpretability of the models, and practicality to the industry.