Energy Forecasting

REQUEST FOR PROPOSAL
RFP #: EF – F2.H1

TITLE: Energy Forecasting
CLOSING DATE AND TIME: OCTOBER 9, 2024 @ 5:00 PM

Hydro-electricity Forecast: EF – F2.H1

Background and Purpose

By responding to this Request for Proposal (RFP), the Proposer agrees that s/he has read and understood all documents within this RFP package.

Submission Details

Responders to this RFP should supply the following 2 files (uploaded to Moodle) by October 9, 2024 @ 5:00 PM:

1. A business report up to 5 pages (not including cover page or table of contents), including any supporting plots and tables.
2. The commented code (in a separate file) used to produce the results.

The report should address **all points described in the "Objective" section** below.

Background

More states are starting to move towards a deregulation of energy companies which allows consumers the ability to choose their own energy supplier. But what does this mean? There are 3 parts of your electric service:

1. Generation – the production of electricity
2. Transmission – the movement of electricity from where it is generated to the point of distribution
3. Distribution – the delivery of electricity to your home or business

When you shop for an electric supplier, you are choosing the company that generates your electricity. Generation supply costs make up most of an average electric bill, so savings may be significant. For most electric customers who select a new supplier, the transmission costs will typically be included in the new charges received from the new supplier. The electric utility that distributes your electricity will remain the same.

Your team is tasked with forecasting the electric load for American Electric Power Co. (AEP), one of the major electricity suppliers in mid-Atlantic. Specifically, you will be forecasting hourly energy load for the Appalachian Power Territory of AEP. Electricity load by a power supplier need forecasts to better meet the needs of customers and prepare for operational expenses.

You are to forecast hourly energy usage for October 25, 2024 – October 31, 2024. Every week you will be sent the updated data with the latest week's energy data through October 24, 2024.

The project will be broken down into 3 phases:

- Phase 1 – Exponential smoothing and seasonal ARIMA models
- Phase 2 – Prophet and neural network models
- Final Phase – Additional model as well as overall comparison of all modeling approaches used.

Objective – Phase 1

The scope of services includes the following:

- The data contains monthly information with an annual seasonal component:
  - Feel free to try different approaches to account for the seasonality.
  - Explain which approach you use and why.
- Build an appropriate Exponential Smoothing Model.
  - Forecast this model for your validation set only.
  - Calculate the MAE and MAPE for the validation set.
- Build a seasonal ARIMA model.
  - Describe the approach you used to select the lags of the model.
  - Forecast this model for your validation set only.
  - Calculate the MAE and MAPE for the validation set.

Data Provided

- The data is provided in a csv format. The variables contained within this data set are:

Datetime_beginning_ept – Date of the observation (in hourly format)
Nerc_region, mkt_region, zone, load_area – unused variables for defining region
Mw – megawatt hour energy usage for the given hour

(HINT: Getting the index to work correctly with hourly data can be a little difficult in the tsibble. One of the ways of doing this is to use the as.POSIXct() function.

```
TRAINING_DATASET_NAME$datetime_beginning_ept <-
as.POSIXct(TRAINING_DATASET_NAME $datetime_beginning_ept,
                                    format = "%m/%d/%y %H:%M",
                                    tz = "America/New_York")
```

This will get you a workable date variable that can be put into a tsibble object as an index. The only other thing to worry about is daylight savings time. You will have to account for multiple 2:00am's in the spring and a lost 2:00am in the fall. Feel free to impute these as you see fit! Welcome to real data ☺)