

Computer Science Department
CS675 – Introduction to Data Science (CRN: 73405)

Fall 2024

Project #3 / Due 03-Dec-2024 (10-Dec-2024)

Implement a Time Series Forecasting model in Python, by using the FBProphet module.

The forecasting model should be able to predict **New York City's Electricity Consumption** (see below) by using Facebook's Prophet model. Prophet is a procedure for forecasting time series data based on an **additive** model where non-linear trends are fit with yearly, weekly, and daily seasonality.

You should test your forecasting model in three (3) distinct datasets. On Daily, Monthly Mean, and Yearly Mean electric consumption.

Daily data:

What's in this Dataset?

Rows	Columns	Each row is a
448K	27	Electric Consumption

Dataset contains daily electric consumption for all five (5) boroughs of New York City.

https://data.cityofnewyork.us/Housing-Development/Electric-Consumption-And-Cost-2010-Feb2023-/jr24-e7cr/about_data

Monthly Mean data:

Take the above (daily dataset) data and average it out based on each month.

Yearly Mean data:

Take the daily data and average it out based on each year.

Write Python scripts in order to complete the following tasks along with their output. All work should be done and submitted in a single Jupyter/Google Notebook.

1) Since the time unit (day, month, year) varies from dataset to dataset, make your code agnostic of the input. In other words, have your code to determine the unit of the time series.

2) Then, train your model (on the respective dataset) and predict the Electric Consumption (EC) values from the last date of the dataset into X units of time into the future.

a) Should the unit of time be day, then predict the EC for 100/200/365 days into the future.

b) Should the unit of time be month, then predict the EC for 1/6/9 months into the future.

c) Should the unit of time be year, then predict the EC for 1/10/20 years into the future.

3) Tune your FBProphet model on the following parameters:

a) **Forecasting growth**: Plausible values = logistic; linear; flat

https://facebook.github.io/prophet/docs/saturating_forecasts.html

b) **Seasonality**: Add manual seasonality by using the add_seasonality method. Test it with various values for 'period' and 'fourier_order'.

https://facebook.github.io/prophet/docs/seasonality_holiday_effects_and_regressors.html

c) **Trend Change points**: Tune the 'n_change points' and 'change point prior scale' arguments/parameters

https://facebook.github.io/prophet/docs/trend_change_points.html

For each model, print the predicted values in a tabular format and draw a line graph showing both historical data and the future.

4) Evaluate all models by providing their respective **MAE** (Mean Absolute Error) and **MAPE** (Mean Absolute Percentage Error), as well as **R²** (use sklearn's respective metrics).

Here are details about the daily dataset (timeseries). You need to manually create the monthly and yearly timeseries.

Daily NYC Electric Consumption: https://data.cityofnewyork.us/Housing-Development/Electric-Consumption-And-Cost-2010-Feb-2023-/jr24-e7cr/data_preview

Electric Consumption And Cost (2010 - Feb 2023)

Search

For Name	UMIS BILL ID	Revenue Month	Service Start Date	Service End Date
/ YORK POWER AUTHORITY	12243668	2023-01	12/23/2022	01/25/2023
/ YORK POWER AUTHORITY	12243668	2023-01	12/23/2022	01/25/2023
/ YORK POWER AUTHORITY	12243668	2023-01	12/23/2022	01/25/2023
/ YORK POWER AUTHORITY	12243668	2023-01	12/23/2022	01/25/2023
/ YORK POWER AUTHORITY	12243668	2023-01	12/23/2022	01/25/2023

Extra Points: Predict Electric Consumption for each of the 5 Boroughs (independently)!