Computer Science Department
CS675 – Introduction to Data Science (CRN: 27962)
Spring 2022

Project #3 / Due 06-May-2022(*)

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

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

Sunspots are temporary phenomena on the Sun's photosphere that appear as spots darker than the surrounding areas. They are regions of reduced surface temperature caused by concentrations of magnetic field flux that inhibit convection. Sunspots usually appear in pairs of opposite magnetic polarity. Their number varies according to the approximately 11-year solar cycle.
Source: https://en.wikipedia.org/wiki/Sunspot

You should test your forecasting model in three (3) distinct datasets. On Daily, Monthly Mean, and Yearly Mean sunspots.
Daily data:
http://www.sidc.be/silso/infosndtot

Monthly Mean data:
http://www.sidc.be/silso/infosnmtot

Yearly Mean data:
http://www.sidc.be/silso/infosnytot

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 Notebook**, or **Python (.py) file.**

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 Sunspots 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 # of sunspots for 100/200/365 days into the future.
b) Should the unit of time be month, then predict the # of sunspots for 1/6/9 months into the future.

c) Should the unit of time be year, then predict the # of sunspots 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 Changepoints: Tune the 'n_changepoints' and 'changepoit_prior_scale' arguments/parameters
https://facebook.github.io/prophet/docs/trend_changepoints.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 R^2 (use sklearn's respective metrics).

Here are details about the daily dataset (timeseries). Find the monthly and yearly timeseries details at the url's provided above.

Daily total sunspot number: http://www.sidc.be/silso/infosndtot

Filename: SN_d_tot_V2.0.csv
---------- CSV ----------
Filename: SN_d_tot_V2.0.csv Format: Comma Separated values (adapted for import in spreadsheets) The separator is the semicolon ';'.
Contents:

- Column 1-3: Gregorian calendar date
  - Year
  - Month
  - Day
- Column 4: Date in fraction of year.
- Column 5: Daily total sunspot number. A value of -1 indicates that no number is available for that day (missing value).
- Column 6: Daily standard deviation of the input sunspot numbers from individual stations. Column 7: Number of observations used to compute the daily value.
- Column 8: Definitive/provisional indicator. '1' indicates that the value is definitive. '0' indicates that the value is still provisional.

Submit **3 different JNBs**, one for daily forecasting, one for monthly, and another for yearly.