Time series Analysis & Modeling

DATS 6450

Homework#4- Simple & Advanced Forecasting Methods

You will need the following libraries to perform the tasks in this homework:

```
import pandas as pd
import matplotlib.pyplot as plt
import statsmodels.tsa.holtwinters as ets
import numpy as np
from pandas.plotting import register_matplotlib_converters
from sklearn.model_selection import train_test_split
```

In this homework, you will use the simple and advanced forecast methods on 5 different dataset and compare the performance of 6 simple and advanced forecasting methods:

1. **Average Method**
2. **Naïve Method**
3. **Drift Method**
4. **Simple Exponential Smoothing Method (alfa=0.5)**
5. **Holt's Linear Method**
6. **Holt-Winter Seasonal Method**

Develop a python program (using above libraries and the codes developed in the previous LABs/HW) to perform the following tasks:

Load the Sales values of dataset "AirPassengers.csv". Split the dataset into training and testing sets (20% test). Use the following command from the 'sklearn' package:

```
train_test_split(y, shuffle= False, test_size=0.2)
```

1- Using all 6 forecast methods above, perform h-step ahead prediction (h is the length of test set). The goal is to compare the h-step ahead prediction versus the test set.
2- Plot the training set, test set and multi-step ahead forecast in one graph. Add appropriate legend, title, x-axis, and y-axis label.
3- Calculate mean square error (MSE) of the forecast error for all 6 forecast methods.
4- Display the variance of prediction error versus the forecast error on the console for each method. Write down your observation.
5- Plot the ACF of the forecast errors for each forecast. (# of lags 15)
6- Calculate the Q value for each forecast.
7- Calculate the correlation coefficient between the forecast errors and the test set. Write down your observation and justify your answer.
8- Create a table and put all above information in the table and conclude which of the 6 forecast method perform the best for this dataset. You need to justify your answer.
9- Repeat set 1 through 10 on the following dataset

- "Shampoo.csv"
- "Daily-total-female-births.csv"

- “Tute1.csv” (Sales needs to be forecasted)
- “daily-min-tempretures.csv”

Be ready to upload the **solution report (as a single pdf**) plus **.py files** through BB.