CS/CE 457/464 - Homework Assignment 9: Time Series Analysis

Due Date: Monday, November 11 at 11:59 pm

Purpose:
Demonstrate understating of Time Series Analysis for prediction and forecasting.

Points: 100

Deliverables: Submit ipynb code file along with your answer

- Use the dataset `karachi-weather-2021-2023.csv`
- Perform time series analysis on the following questions**.** Make sure to include interpretation of each result including visualizations to support your answer

1. Aggregate the data into weekly mean. Use `.resample('W').mean()`
2. Read the dataset visualize the trend, seasonality, residual etc. Discuss your observations in the visualization.
3. Split the data into train and test sets. Use all the data for 2021-2022 for training and 2023 for testing.
4. Use ARIMA, visualize the results and report the RMSE.
5. Use SARIMA, visualize the results and report the RMSE. Use '52' as the fourth parameter of seasonal_order (for example `seasonal_order=(0, 0, 0, 52)`) since there are 52 weeks in a year.
6. Try to come up with optimal parameters for SARIMA. Pick the best one, visualize the results and then report the RMSE. Compare both ARIMA and SARIMA and conclude which one is better?
7. Now use any Machine Learning Regression technique and use the same training and testing set (used for ARIMA and SARIMA). Visualize the results and report the RMSE.
8. Compare all the RMSEs (ARIMA, SARIMA and ML Regression algorithms) and discuss which one is better.
9. Generate future dates (weekly) from Nov 1st 2024 to December 31th 2024 and predict the price using SARIMA model (with optimal parameters). Use `pd.date_range(start='01/11/2024', end='01/12/2024', freq='W')`
10. Using the same future dates, predict the price using any one ML regression model
11. Visualize and analyze the predictions from 9 and 10 and discuss their differences.

Additional Task for Learning (ungraded)

1. Use exogenous variables of your choice from `karachi-weather-2021-2023-more-variables.csv` in SARIMA and evaluate if it helps improving the prediction of a model.