

Analysing sensor data

Lab 3

Background:

Sensors produce autonomous data which can be collected continuously or periodically. Sensors work like a minicomputer with its own set of operations. Working with sensor data will help to automate actions.

Purpose:

The purpose of this assignment is to learn how to analyse and forecast sensor data. Sensors produce data which can be collected over a period. This historical data can be utilized to analyse and perform predictions.

ARIMA:

Autoregressive (AR): This component uses lagged values of the time series to predict future values. The number of lagged values used is represented by the parameter 'p' in the ARIMA model.

Integrated (I): This part involves differencing the time series to achieve stationarity. The degree of differencing is denoted by 'd'. A stationary time series is one whose statistical properties like mean and variance do not change over time.

Moving Average (MA): The MA component models the error of the model as a combination of previous error terms. The parameter 'q' represents the size of the moving average window.

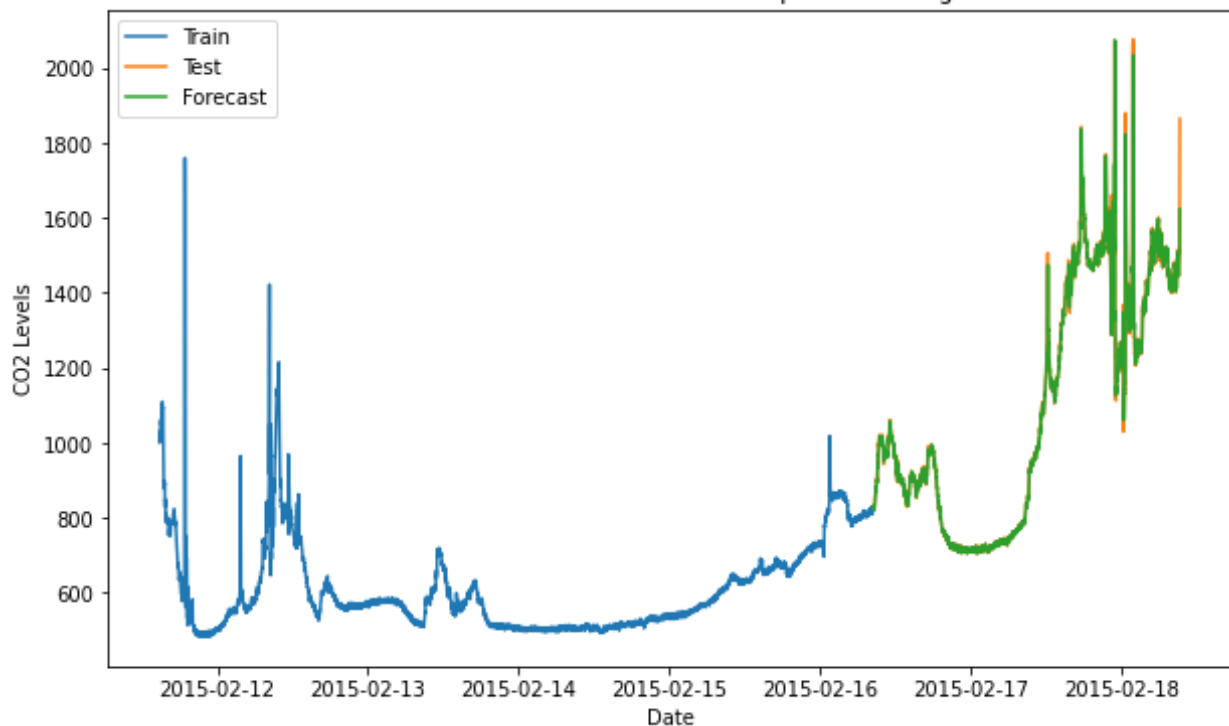
Tasks:

1. You need to analyse multiple columns of the provided dataset. Show scatterplots and correlation matrices of chosen columns. Explain what you understood.
2. Use “adfuller” to understand if your data is stationary
3. Use “auto_arima” to find the best ARIMA model parameters
4. Forecasting: Train and test CO2 data with 70% training and 30% test. You can increase the training size upto 80%.
5. Plot your forecast with actual values as shown in the below figure. Use stepwise approach for better forecast values.
6. Show MAE value of the test
7. Discuss your results and provide your reflection

Presentation: Record a video in which you explain your results and upload it to your YouTube channel. Maximum time of recording should be 10 minutes.

Lab 3 needs to be submitted no later than **Jan/10 23.59** in the submission folder on Learn/Assignment/Labs/Lab3.

Your target should be something like this:



HINT: Use Step wise training approach while testing

1. Initial Training: Train the ARIMA model on the initial training set.
2. Stepwise Update and Forecast:
 - For each data point in the test set, forecast the next step.
 - Update the model with the actual observed value.
 - Store the forecast and the true value for later evaluation.

Material: All material for the tasks can be found in Lab2 folder in Learn/Assignments/Labs.

Grade:

This assignment is graded as U/G.

G will be given if all the tasks are completed successfully.

U will be given if any task is not completed

Note: Everyone in the group should talk during the presentation.

Reference:

How to interpret ADFULLER test

```

(-5.717572244272956,
7.062691902377055e-07,
19,
559,
{'1%': -3.442102384299813,
'5%': -2.8667242618524233,
'10%': -2.569531046591633},
-617.2404579285969)
# The test statistic.
# p-value
# Number of lags used
# Number of observations used
# Critical values (1%, 5%, 10%)
# The maximized information criterion

```

- ✓ Stationary if $p\text{-value} \leq 0.05$
- ✓ Stationary if $|crit.\text{value}| < |test\ statistic|$