Stock Forecasting

Rupali Shah | Springboard | Sept 28, 2019

Overview

This part of the project uses the Inflation Rate and Adjusted Closing price to forecast stock prices.

Vector Auto Regression (VAR)

VAR is one of the most commonly used methods for multivariate forecasting. In a VAR model, each variable is modeled as a linear combination of past values of itself, and the past values of all the other variables in the system. The VAR model can understand the relationship between several variables.

Data Munging

The column names of the Inflation Rate dataset are renamed to make them consistent with the columns of the Stock Price dataset.

The Inflation Rate column is converted to numeric.

The missing values are filled by averaging the values of the previous and the next day.

The Stock Price and the Inflation Rate datasets are merged.

Stationarity Check

The multivariate series is checked for stationarity based on the Eigen values. This is done using the Coint Johansen test.

The Time Series is on a 'Business Day' Frequency, as there is no trading on weekends. The gap in the series for Saturday and Sunday is filled by first converting it to a Daily Time Series, and then filling the values of stock prices for weekends using a Forward Fill.

Model Building

The dataset is split into train and test while accounting for the time component. The train_test_split and k-fold validation cannot be used here, since it can disrupt the pattern in the Time Series. The data is then fit to the VAR model letting the model determine the number of lags. Predictions are made on the test dataset.

The model gave an RMSE of 0.0127