Table of Contents

Stock Prediction ARIMA vs LSTM

Introduction

One of the ways to ensure higher returns than saving in the back is by investing in the stock market. Large financial institutions as well as individuals would like to maximize returns on their money, and one of the substantial ways this can be achieved is through the investing in the stock market. There are various indicators and tools that are available to help financial investors and traders make calculated decisions on what stocks to buy but these indicators but even with these it has been a very hard task to predict the future prices of stocks. As technology has advanced the financial industry has been interested in using machine learning in price prediction, pricing and managing entire portfolio of assists and investment process. With increase in computational efficiency and deep learning the applications are very versatile and we would like to explore the use of these new tools in stock prediction. With recent combination of machine learning models and statistics algorithms have been made that can reveal intricate patterns that are non-linear.

In this project we will look into a statistical based prediction algorithm ARIMA in particular and compare it to the the technologically advanced neural network deep learning system LSTM RNN in particular and see how it performs and if there is true advancement between the traditional and advanced method of stock prediction.

Objective

The main objective of this project is to access and investigate with forecasting model produce the best prediction results, by analyzing the model with lower forecast errors using mean square error and mean average error. With modern technology deep learning is able to patterns and structure in data that are more complex this project seeks to investigate its effectiveness in stock prediction.

Data

Standard stock price data has five columns the high, low, open, close, adjusted close and date. These dates do not include the weekend and holidays. The data ranges form 2012-01-05 to 2019-12-31. About a total of 8 years, making a total of 2010 rows with 6 columns. The high signifies the maximum price for the day while the low indicates the minimum price for the day, the open is the price the stock was at the beginning of the day, the close is the price at the end of the day, the adjusted is the price at the end of the day but factors in anything that might affect the stock price after the market closes.

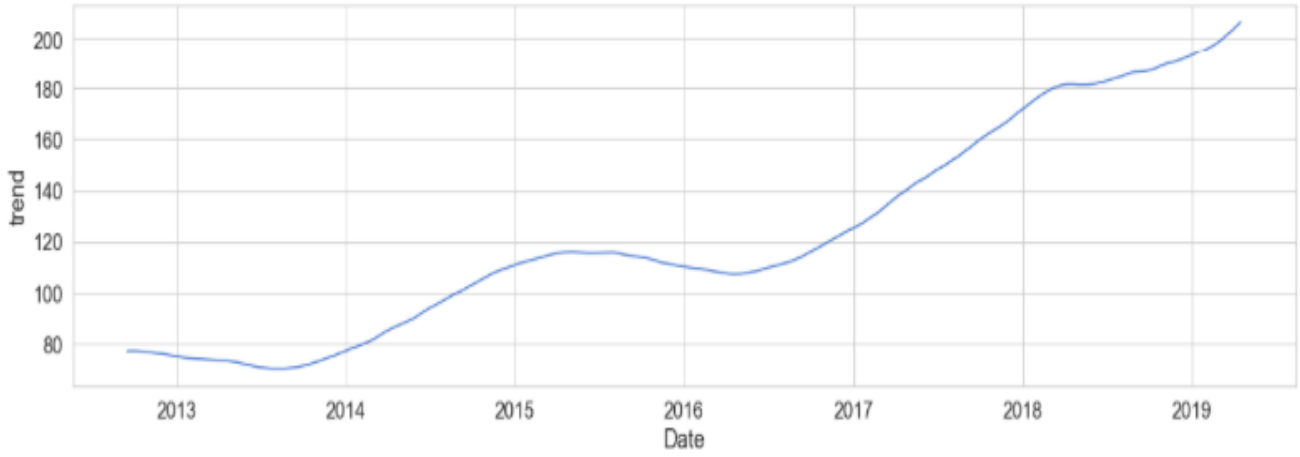
Figure2: close price 2012 - 2019

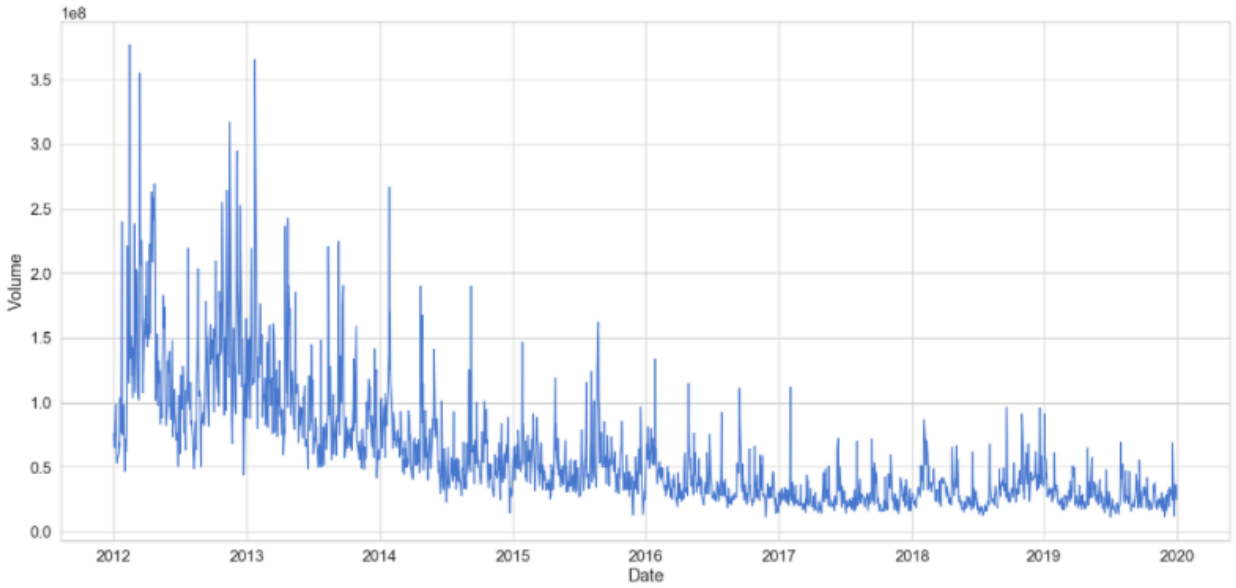
Data Exploration

Data exploration with regards to stock charts is not broad as the main technique is to use past data to predict the present. The only feature used in this study would be the daily closing costs of the stock.

1. Decomposition of Time Series

A series needs to be decomposed in order to ensure that there are some patterns and that the series is not majorly white noise which cannot be used for or casting.





a. Seasonal Chart from this chart we can see a cyclic pattern that occurs but over the a period of a year.
b. Trend: shows the overall direction of the apple stock price over the course of years
c. Residuals this is the unexplained noise that is found in the data, over those number of years after you remove the effect of the trend and the

3. Check for Patterns over smaller time range
    a. In order to look more closely at data on closer basis we check if we can find any pattern in the time series over the span of a month. From this chart there is no constant pattern that occurs monthly.
    b. This chart tells us high high the buying and selling interests are , here we see volume of shares traded over the years has reduced considerably.