

AAPL analysis Report

Prepared by:

Srija Uprety
Bachelors (Sixth Semester)

1. Objective and Expected Significance

The major objective of this project was to improve the deep learning models (via hyperparameter optimization) that may be effective in solving prediction issues like stock market value.

It must be made clear that my intention was not to utilize this model to predict the stock market for financial gain, but rather to use it as an opportunity to comprehend and gain knowledge used in hyper parameter optimization of complex predictive models.

2. Data Source

I have limited our project to predict the closing price of AAPL Stock. The data has been extracted from publicly accessible platform Yahoo. Data from July 31, 2017 to July 29, 2022 were taken into account. The data consists of Opening value, High and Low value, closing value, adjusted closing value and the volume for Apple Stocks. We are concerned about predicting the adjusted closing value. There are 1259 instances i.e. the number of days.

3. Approach

I have contemplated two models widely used in Time Series Forecasting. One of the models is LSTM(Long-Short-Term-Memory) and the other one is ARIMA(Autoregressive Integrated Moving Average) which also acts as a baseline for its effectiveness in widespread time series analysis. I have specifically focused on closing data of AAPL stock data of the past 5 years.

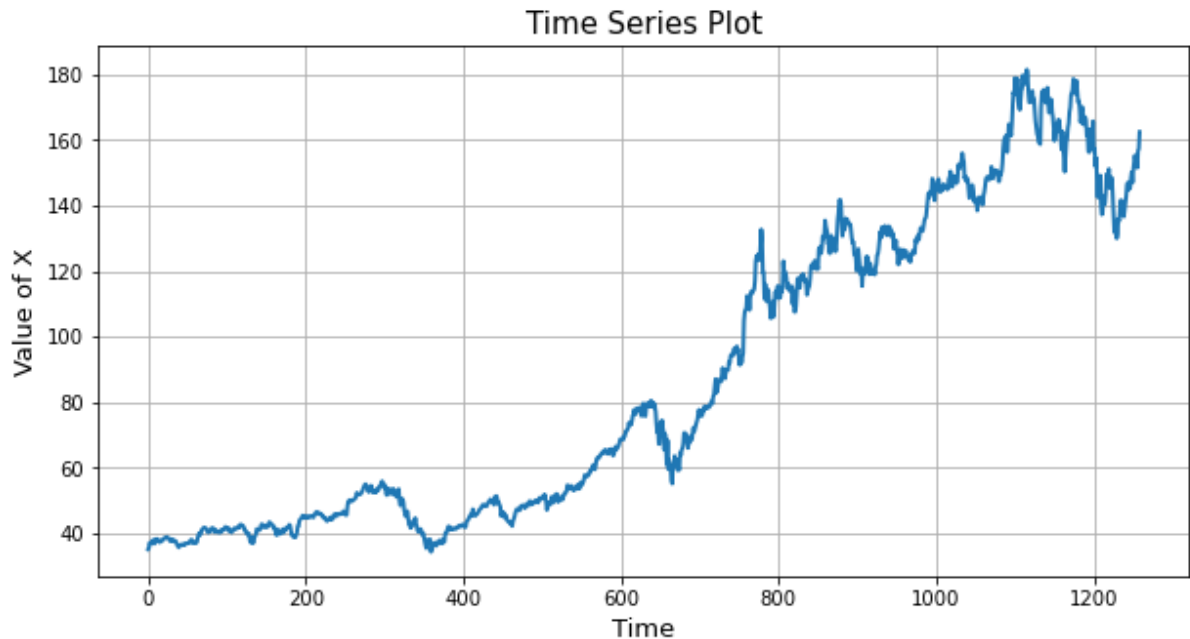


Figure 1: Adjusted Closing Value of AAPL

This is the Adjusted closing value of AAPL. I will briefly explain about the models that I used for predicting the closing values. The value of X determines the closing value of the data.

4. LSTM model

It is a network that operates on the current input while taking into account the prior output (feedback) and temporarily storing it in memory (short-term memory). It stores the output for a short period of time. The network has a visible layer with 1 input, Three hidden layers with 40 LSTM blocks or neurons, and an output layer that makes a single value prediction. The default sigmoid activation function is used for the LSTM blocks. The network is trained for 200 epochs and a batch size of 32 is used.

Layer (type)	Output Shape	Param #
layer1 (LSTM)	(None, 7, 40)	6720
layer2 (LSTM)	(None, 7, 40)	12960
layer3 (LSTM)	(None, 40)	12960
outputlayer (Dense)	(None, 1)	41
=====		
Total params: 32,681		
Trainable params: 32,681		
Non-trainable params: 0		
=====		
None		

Figure 2: LSTM Network

5. ARIMA(Automated Regressive Integrated Moving Average) model

Based on its own prior values, the ARIMA model makes predictions for a given time series. ARIMA model only works when the data is stationary. The model followed three steps for its implementation.

- The data was made stationary.

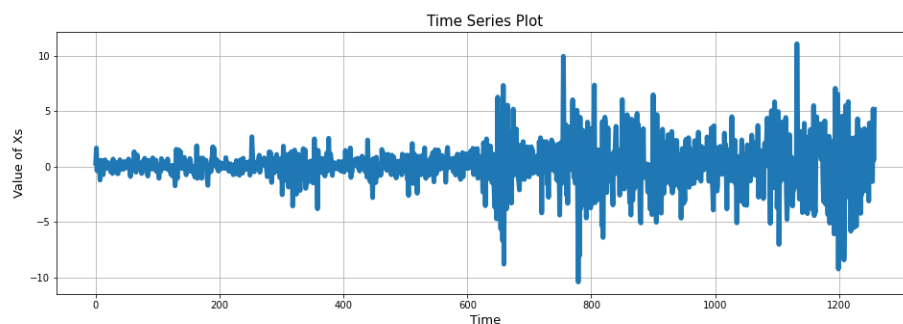


Figure 3: Trend data turned into stationary data

- Estimated ACF and PACF values. Observing the ACF plot, the ACF values appear to cut off after lag 0. So we set \hat{q} to 0. Similarly, from the PACF plot, the PACF values appear to cut off after lag 2, so we set \hat{p} to 2.

- Fitted ARIMA model with order (p,d,q) where d is the differentiating order.

ARIMA Model Results						
Dep. Variable:	D.y	No. Observations:	1258			
Model:	ARIMA(2, 1, 0)	Log Likelihood	-2640.510			
Method:	css-mle	S.D. of innovations	1.974			
Date:	Sat, 30 Jul 2022	AIC	5289.020			
Time:	23:40:22	BIC	5309.569			
Sample:	1	HQIC	5296.743			
	coef	std err	z	P> z	[0.025	0.975]
const	0.1009	0.052	1.943	0.052	-0.001	0.203
ar.L1.D.y	-0.0615	0.028	-2.176	0.030	-0.117	-0.006
ar.L2.D.y	-0.0101	0.028	-0.357	0.721	-0.065	0.045
Roots						
	Real	Imaginary	Modulus	Frequency		
AR.1	-3.0468	-9.4769j	9.9546	-0.2995		
AR.2	-3.0468	+9.4769j	9.9546	0.2995		

Figure 4: ARIMA Network

6. Project Outcome

Figure 5 and Figure 6 summarizes the outcome of LSTM and ARIMA model respectively.

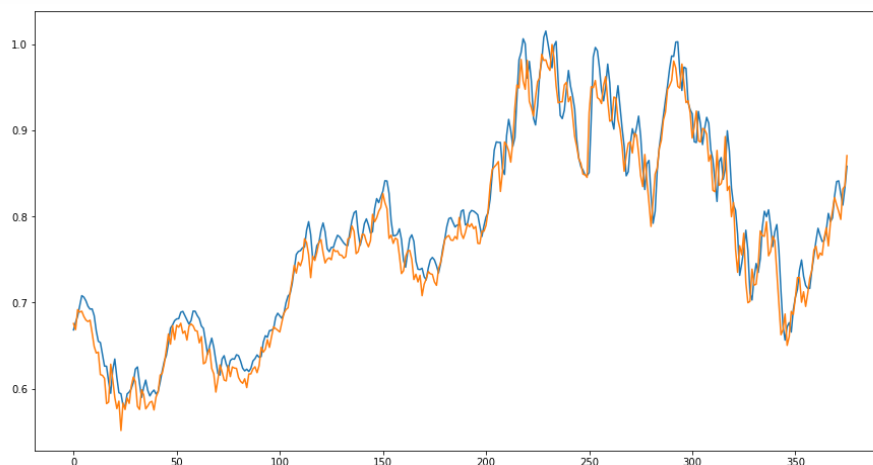


Figure 5: Predicted and Observed data of Closing value of AAPL using LSTM

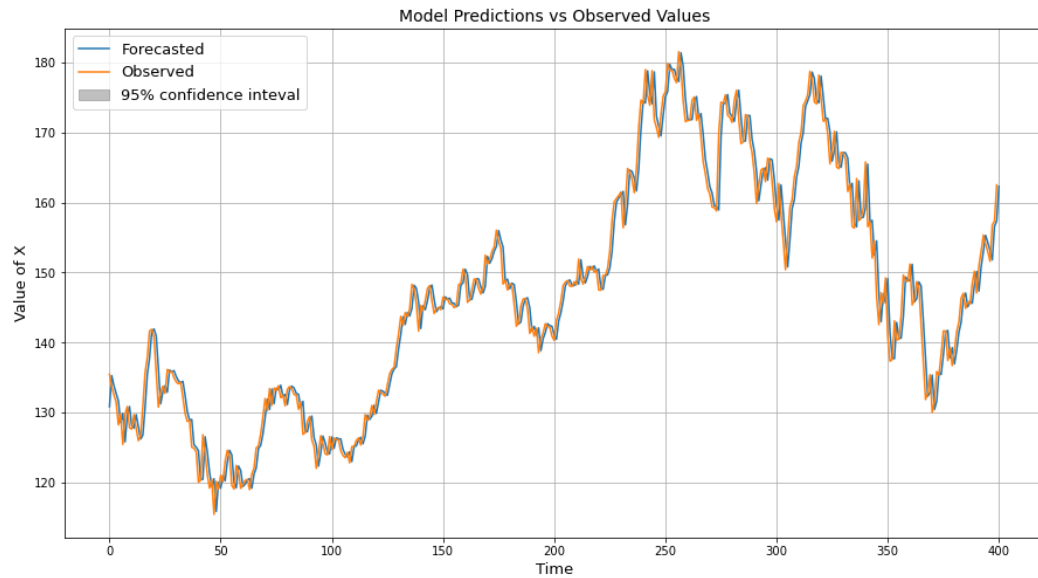


Figure 6: Predicted and Observed data of Closing value of AAPL using ARIMA

The following conclusions were made based on the experiments:

- Epochs and batch sizes are a major hyper parameter to consider if the computational cost needs to be taken into account.
- ARIMA model had better accuracy than LSTM and it could be widely used to predict the forecast as it uses lagged moving averages to smooth time series data.

7. References

[1]Time Series Prediction with LSTM Recurrent Neural Networks in Python with Keras. In Machine Learning Mastery. Available from:

[Time Series Prediction with LSTM Recurrent Neural Networks in Python with Keras \(machinelearningmastery.com\)](https://machinelearningmastery.com/time-series-prediction-lstm-recurrent-neural-networks-python-keras/)

[2]Yahoo Finance:NasdaqGS - NasdaqGS Real Time Price. Currency in USD

<https://finance.yahoo.com/quote/AAPL/history?p=AAPL>

[3]ARIMA Documentation

<https://www.statsmodels.org/devel/generated/statsmodels.tsa.arima.model.ARIMA.html>