

Report on Stock Prediction Model

-Ananya Rai

I. MOTIVATION

Design and implement a machine learning model for predicting the future value of stocks based on historical price data and market indicators. The model should be capable of generating accurate predictions over various time zones.

II. DATASET DETAILS

We used Yahoo Finance's dataset and the respective YFinance python library to build the dataset for this stock prediction model. Yahoo Finance hosts the data for every major listing on NASDAQ(National Association of Securities Dealers Automated Quotations) and provides the open price, the high, the low, and the close price of every single trade day for that listing since the documented history and to the present date.

We intend on using the close price variable to predict the values, thus making our model univariate. For the demonstration, we have run our model on the stock history of Apple and Microsoft and their stock trends from 1st January 2016 to 1st January 2024. We have divided the dataset into two parts - namely train(65%) and test(35%).

III. PROPOSED ARCHITECTURE

We intend to use Long Short Term Memory Neural Networks(LSTM) for our architecture - majorly because of its ability to learn from sequential data while considering long-term dependencies. In the context of stock prediction, past stock prices and other relevant factors form a time series, and LSTM can effectively capture patterns and trends within this sequential data. Also their ability to handle non-linear data and memory retention, they seem to be the perfect model for our application.

We have used TensorFlow's Sequential model with LSTM and Dense layers for libraries. The model trains and predicts for each day using the trends of the last 100 days to make its decisions. We have used 3 layers for LSTM and one layer for Dense, and have computed our losses in terms of mean-squared error.

IV. RESULTS

I) *Apple*:

- MSE on the train and test set at the last epoch of the LSTM model for each data point (mean) were $9.9508e-05$ and 0.0019 respectively.
- The net MSE on train set and test set were found out to be 62.112 and 154.387 respectively.
- The future prediction seem a little flawed considering the huge dips and rise in our predictions compared to the past trends.

II) *Microsoft*:

- MSE on the train and test set at the last epoch of the LSTM model for each data point (mean) were $9.5480e-05$ and 0.0012 respectively.
- The net MSE on train set and test set were found out to be 125.20580 and 284.346768 respectively.
- The future prediction seem a little flawed considering the huge dips and rise in our predictions compared to the past trends.

V. ANALYSIS OF RESULT

- LSTM proved to be a good regressor for something like stock prediction which is factored by a lot of real world variables which one might say to be random, but using the past data it was able to give a good estimate.
- Our model fails to predict for large durations in the future because of its data being referenced from another future predicted data. At best, it can give a prediction for the next day or two.

VI. REFERENCES

- Greg Hogg's youtube video on LSTMs and Stock prediction model.
- Krish Naik's youtube video on LSTMs and Stock prediction model.

VII. VISUALIZATIONS

I) *Apple*

The below are our findings when running the model on the Apple's stock history:

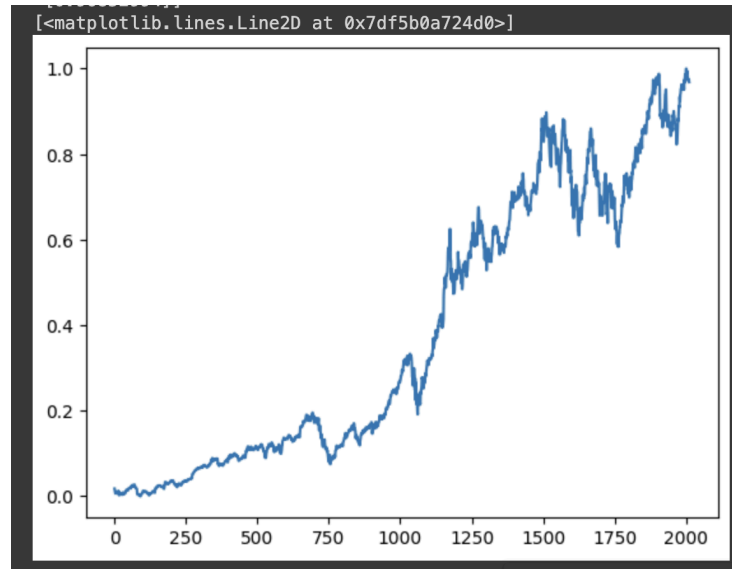


Fig. 1. Apple Dataset

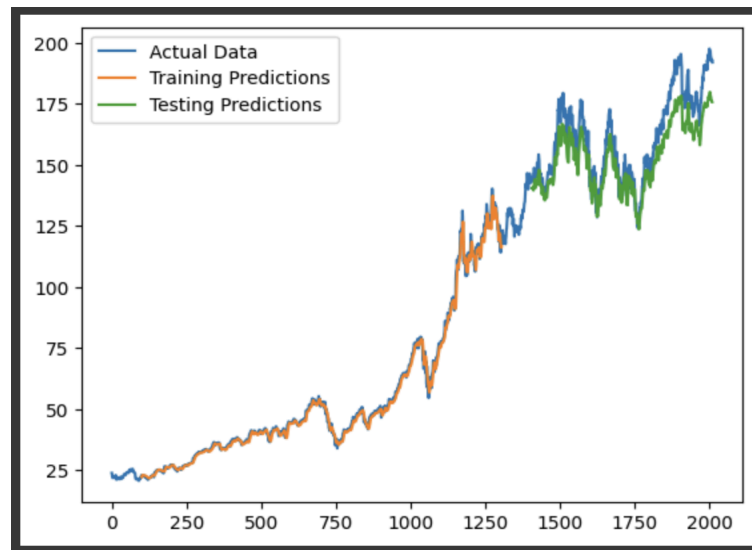


Fig. 2. Predictions on Train and Test set on the Apple dataset

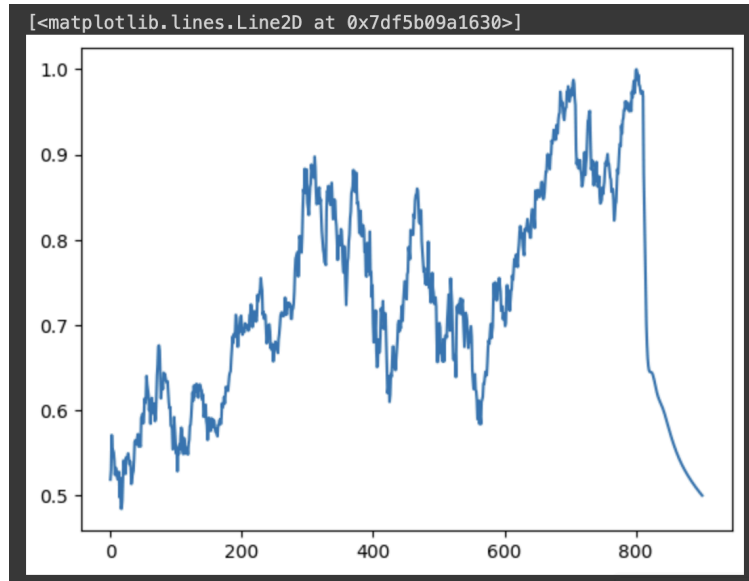


Fig. 3. Apple predictions for next 90 days

II) *Microsoft*

The below are our findings when running the model on the Microsoft's stock history:

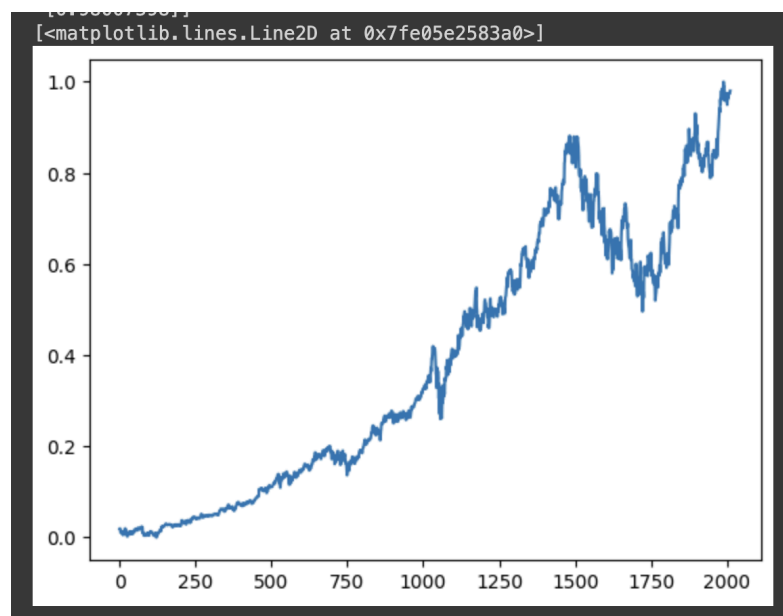


Fig. 4. Microsoft Dataset

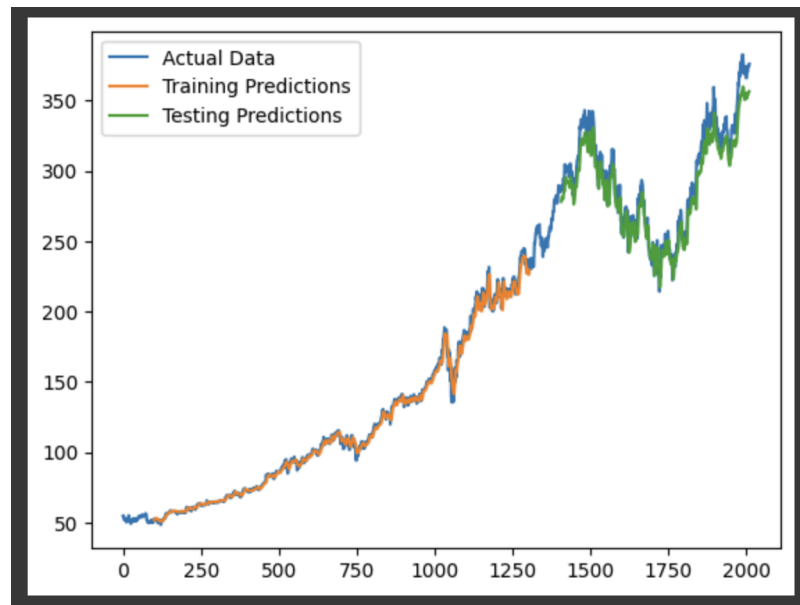


Fig. 5. Predictions on Train and Test set on the Microsoft dataset

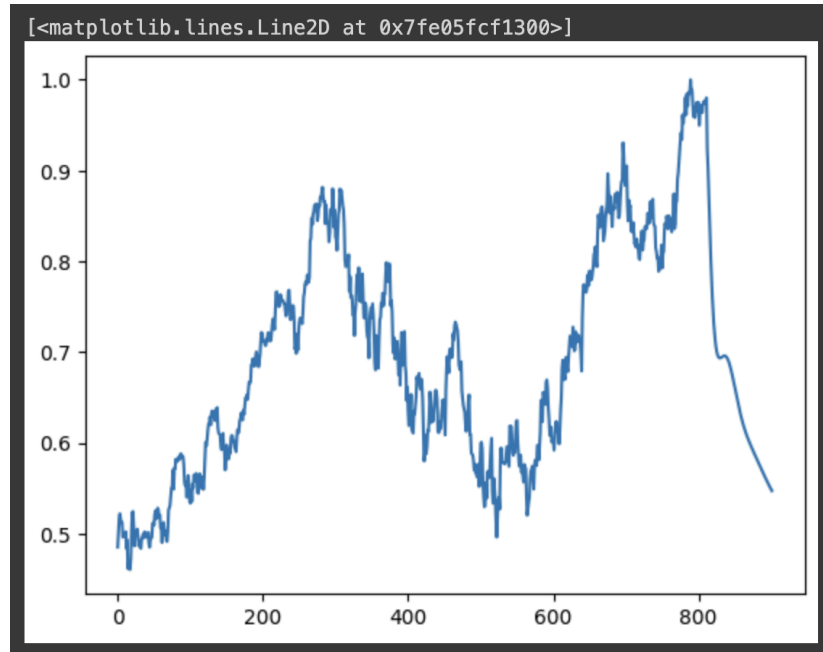


Fig. 6. Microsoft predictions for next 90 days

