Deep Learning based Stock Price prediction

Malipalema Khang
1613006
School of Computer Science and Applied Mathematics
University of the Witwatersrand

***Abstract*—The stock market is a virtual market place where investors buy and sell shares with the intent of making a profit. Accurately predicting the stock price of assets has historically proven to be a tedious and inaccurate process. However, the advancement of technology has led to easier prediction of the seemingly unpredictable stock price movements. This study aims to build an combined model of Long-Short Term Memory model (LSTM) and Gated Recurrent Unit (GRU) neural networks to predict future stock market price. The aim of this study to implement both LSTM and GRU models and find the best algorithm and model parameters to best predict the stock price.**

I. Introduction

Determining the future movement of the stock price is a very profitable aspect of in financial market. There is a constant rise in the meta data of stocks available on financial markets, therefore there is a need to. Research shows that LSTMs followed by GRUs are some of the best models for performing this kind of time series predictions. A study on google stock data [1], studied the effectiveness of Recurrent Neural Networks, and their classes of Long Short Term Memory (LSTM) and Gated Recurrent Unit (GRU), and revealed that LSTMs were the top performing model when compared to standard machine learning models.

This study will make use of both LSTM and GRU models. The LSTM is a class of recurrent neural network architecture excellent for modeling temporal sequences; thus sequences that are highly influenced by the order in which the data is presented in. Time series data is an excellent example of temporal sequences, as the order of the information is highly influenced by the previous data.

GRU is another class of RNNs similar to LSTMs however one advantage of the GRU is that it is less computationally expensive that the LSTM[2]. However the advantage of LSTM over the GRU is has a higher accuracy. This is primarily what motivates this study to build a model based promising the advantage of both speed and accuracy.

II. Background

A. Long Short Term Memory

Long Short-Term Memory (LSTM) is class of Recurrent Neural Networks (RNN), with an added advatage of capturing data from prior stages and recycle it for future predicition [3]. Neural networks mostly consists of three layers:

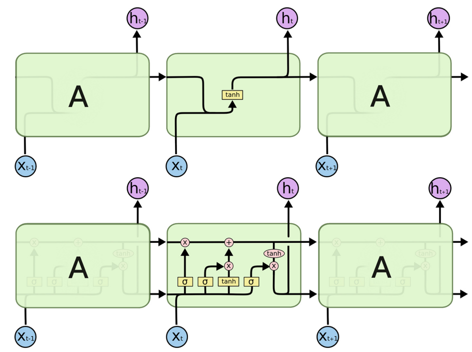

Figure 1. LSTM Stucture

1) input layer
2) Hidden layers
3) output layer

The classes of Neural networks that perform predictive analysis are Recurrent Neural Networks . Just as the name ’recurrent’ suggests, thet used to the make prediction based on observed sequences from prior stages of learning and further use that information to forecast future values. In the instance of this study, the LSTM will make predictions based on previous stock price data. RNNs have a limited architecture since they can’t store long-term memories. The RNN’s apparent limitation was solved with the advent of the Long Short-Term Memory. They have the ability to’memorize’ previous data sequences, making them extremely useful for predictive analysis.The following figure 1 describes the structure of the LSTM.

The ability to memorize data sequences distinguishes LSTMs from other types of RNNs. Every LSTM node must consist of a set of cells responsible for storing passed data streams; the upper line in each cell connects the models as a transport line handing over data from the past to the present; and cell independence aids the model in disposing of a filter of add values from one cell to another.

- The Forget Gate returns a number between 0 and 1, with 1 indicating "completely keep this" and 0 indicating "completely ignore this." - Memory Gate determines which new data is stored in the cell. First, a sigmoid layer called the "input door layer" determines which values will be changed. Following that, a tanh layer generates a vector of new candidate values

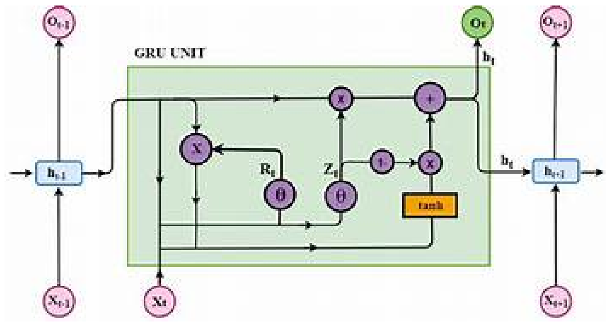

Figure 2. GRU Stucture

that could be added to the state.- The - Output Gate determines what each cell's output will be. The output value will be based on the cell state as well as the most recent filtered and added data.

B. Gated Recurrent Unit

GRU is a type of recurrent neural network that tackles the problem of vanishing gradients. According to [4], a GRU is similar to an LSTM but has fewer gates, as illustrated in 2. It consists of two gates: an update gate and a reset gate. These two gates operate together to govern the flow of data through the network. The update gate determines how much information from the past needs to be sent on to the next step. The amount of data to be forgotten is determined by the reset gate.

III. Implementation

A. Datasets

Apple stock (AAPL) data used for this study was scraped from Yahoo Finance, a global portal that provides analysis and news about global financial markets, as seen in the figure 3. The was collected from over a period of 10 year. From January 1st 2010 to present day, specifically June 15th 2021.

Data	Description
Open	Starting daily of Apple share price
Close	Final daily of Apple share price
High	Maximum daily of Apple share price
Low	Minimum daily of Apple share price
Volume	Volume daily of Apple number of shares bought

B. Model development

To predict the profit or loss for the day will be calculated using the closing price of a stock for the day, that will be used to build the preditive model. Sci-kit Learn's MinMaxScaler was used to normalizing Data: Normalizing data aids the algorithm in converging, i.e. efficiently locating local/global minimum.The train and test data was splint into a 70: 30 split respectively The test data will determine whether our model perfectly validates our data. The data was trained on different model and algorithm parameters, inluding different epochs, activation functions and learning rates, in oder to tune the best parameters for the prediction. All models we tested on

Figure 3. Dataset Visualization

mean squared error metric through Adam's optimizer to get the RMSE values for the predictions.

IV. Research results

A. Research results

The Following table demonstrate the performance of the algorithm on different parameters.

The Following pictures demonstrate the performance of the algorithm on different parameters. The blue represents the actual data and the red represents the predicted.

Model	Train MSE	Test MSE
LSTM 0.05 lr	0.00002	0.02404
LSTM 0.005 lr	0.00001	0.00250
LSTM 0.0005 lr	0.00001	0.00704
LSTM 25 epochs	0.00001	0.00704
LSTM 50 epochs	0.00001	0.00442
LSTM 100 epochs	0.00002	0.02404
LSTM tanh + relu	0.00001	0.00127
LSTM tanh + leakyRelu	0.00001	0.00451
GRU tanh + relu	0.00001	0.00022
GRU tanh + leakyRelu	0.00001	0.00044
LSTM + GRU	0.00001	0.00021

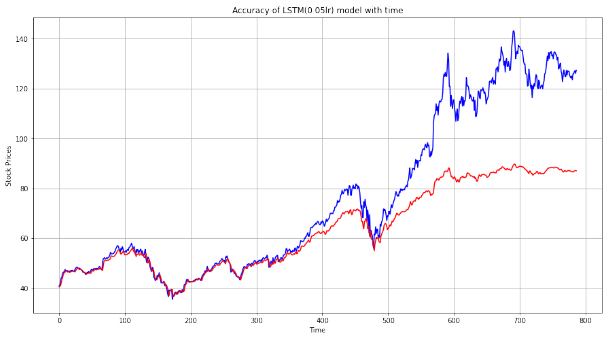

Figure 4. LSTM with learning rate = 0.05

V. Conclusion

The lowest RMSE indicates the best performing model, this was indicated by the LSTM and GRU combined model. The results also demonstrate that adding activation functions, increasing the number of epochs and decreeaseing the learning rate further improve the performance of the model. Overall LSTM and GRU models perform relatively well in stock price prediction. The model can be improved training on different stock datasets as well as including social media sentiment analysis to further improve the performance.

For more details about the implementation please visit: https://github.com/malipalema/Stock-Price-Prediction-LSTM-GRU

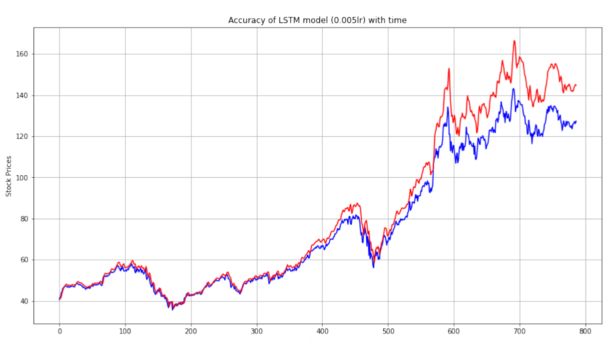

Figure 5. LSTM with learning rate = 0.005

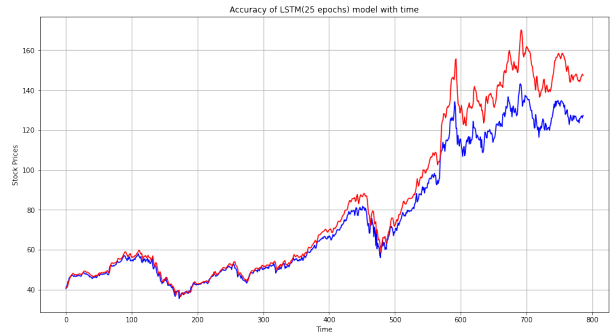

Figure 6. LSTM with epochs = 25

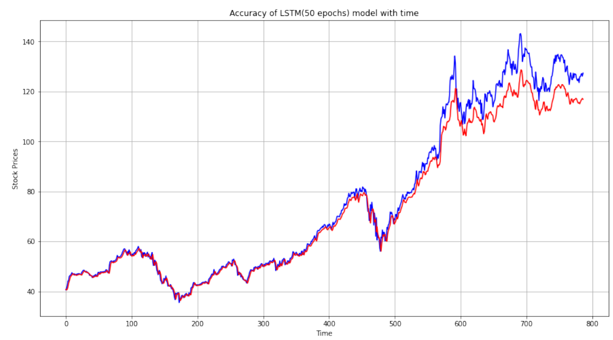

Figure 7. LSTM with epochs = 50

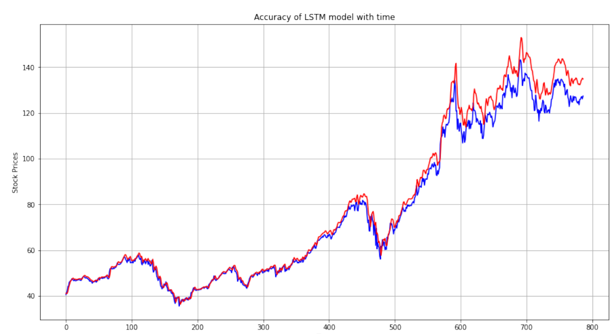

Figure 8. LSTM with epochs 100

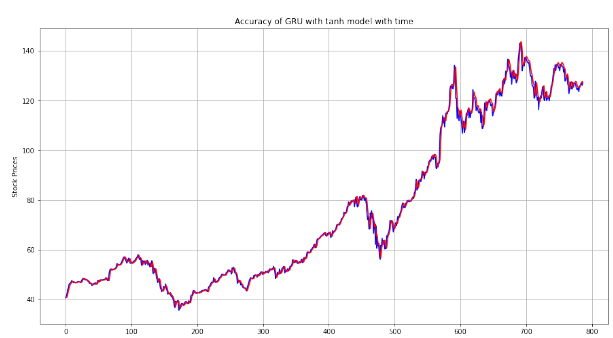

Figure 9. LSTM with tanh

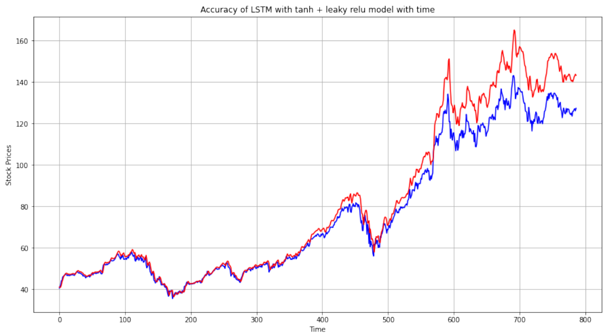

Figure 10. LSTM with tanh + leaky

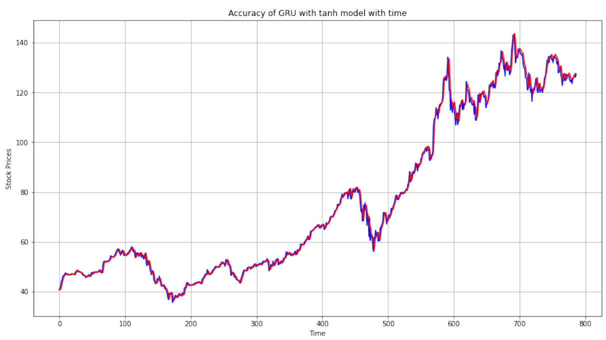

Figure 11. GRU with tanh

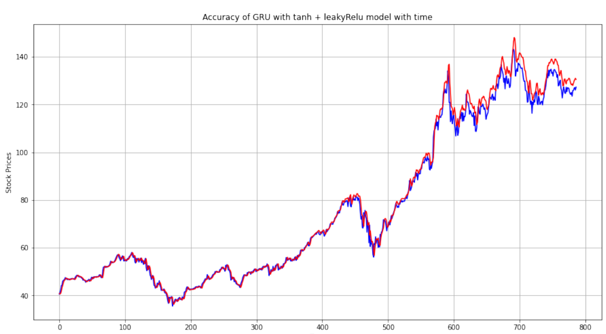

Figure 12. GRU with tanh + leaky

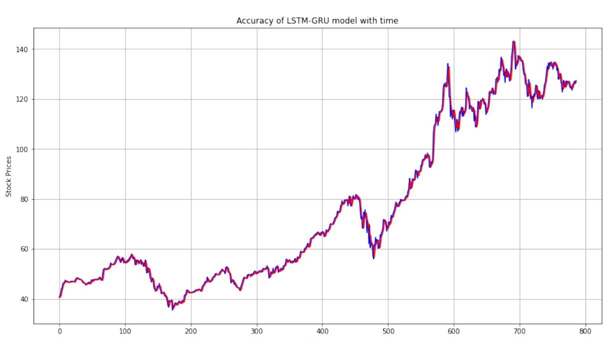

Figure 13. LSTM with GRU

References

[1] Mohil Maheshkumar Patel, Sudeep Tanwar, Rajesh Gupta, and Neeraj Kumar. A deep learning-based cryptocurrency price prediction scheme for financial institutions. *Journal of Information Security and Applications*, 55:102583, 2020.

[2] Md. Saiful Islam and Emam Hossain. Foreign exchange currency rate prediction using a gru-lstm hybrid network. *Soft Computing Letters*, page 100009, 2020.

[3] Adil Moghar and Mhamed Hamiche. Stock market prediction using lstm recurrent neural network. *Procedia Computer Science*, 170:1168–1173, 2020. The 11th International Conference on Ambient Systems, Networks and Technologies (ANT) / The 3rd International Conference on Emerging Data and Industry 4.0 (EDI40) / Affiliated Workshops.

[4] K.E. ArunKumar, Dinesh V. Kalaga, Ch. Mohan Sai Kumar, Masahiro Kawaji, and Timothy M Brenza. Forecasting of covid-19 using deep layer recurrent neural networks (rnns) with gated recurrent units (grus) and long short-term memory (lstm) cells. *Chaos, Solitons & Fractals*, 146:110861, 2021.