Stanford CS230 - Deep Learning Course – Spring 2021

Sadegh Aryanpanah (saryan@stanford.edu)

Course Project Final Report

Prediction of Stock Prices Using Long Short Term Memory (LSTM) and Gated Recurrent Unit (GRU) Recurrent Neural Networks (RNN)

1. Introduction

The stock market is known for its extreme complexity and volatility, and people are always looking for an accurate and effective way to predict future stock prices. Different efforts have been done to use different types of deep learning algorithms for prediction of stock prices. Heaton et al showed that applying deep learning methods to stock price prediction problems can produce more useful results than standard methods in finance.[1]. Among different methods, long short-term memory (LSTM) based recurrent neural networks (RNN) have received high attention in stock price prediction applications. Moghar et al [2] used LSTM RNN to predict the price of some of SP500 stocks in New York Stock Exchange. Akita et al [3] modeled the effect of stock related news in different Japanese newspapers on the opening price of different stocks in Tokyo stock market using RNN LSTM method. Cheng et al [4] reviewed the application of attention in the predication of future stock price RNN LSTM method. Hao et al [5] further increased the accuracy of LSTM RNN methods by considering attention and stacked LSTM model. They used their approach for stock portfolio optimization and maximization of stock profit. Sang et al [6] used both RNN and CNN models in stock prediction and concluded that a RNN LSTM based model is best suited for prediction of stock price because of time dependence of stock price data. In addition to LSTM architecture, gated recurrent unit (GRU) unit has also been used in stock market prediction. G. Shen Et al used GRU for prediction of stock price in different stock market indexes include S&P 500 [7]. M.S. Islam et all used hybrid LSTM and GRU models for FOREX market currency price prediction [8]. In current project, both RNN-LSTM and RNN-GRU models have been constructed, trained and tested for prediction of Google stock closing price. RSME metric has been used to access the performance of both models. The comparison between the accuracies of both models based on RSME metric has been done and the results were reported and disused.

2. Model Description

LSTM model was introduced in 1997 to address the problem of vanishing gradient in RNN model. In 2014, GRU model which is a simpler version of LSTM model with fewer parameters was introduced to address the same problem of vanishing gradient. The illustration and related equations for both GRU and LSTM models are shown in figure 1.

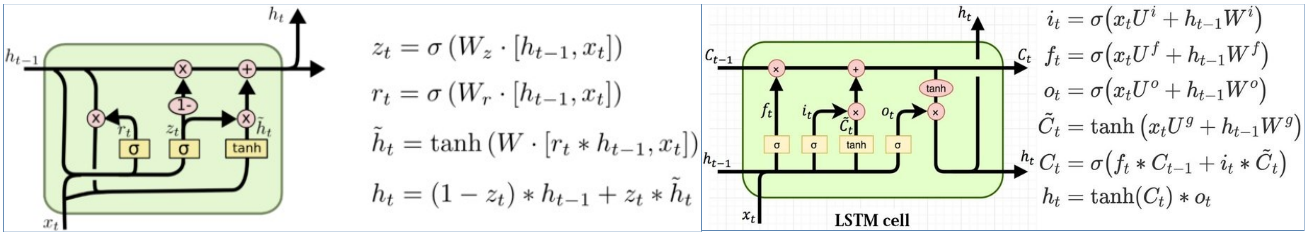

Figure 1. Description GRU (left) and LSTM (right) models

In current project, 2 layers stacked LSTM and GRU models were used for training and prediction of stock closing price data. The intention of current project was the evaluation of the effect of different architectures (LSTM and GRU) on the training and testing residual errors of the stock closing price prediction models. In order to study the effect of architectures on the model outcome, the model hyper parameters for both GRU and LSTM architectures are considered to be same. The model hyper-parameters are mentioned in table1.

Table1. RNN LSTM and GRU Model parameters that were used in current project

GRU and LSTM Parameters	Value
Number of Stack Layers	2
Number of Features in Input Layer	1
Number of Features in Output Layer	1
Number of Features in Hidden State	32
Learning Rate	0.01
Loss Function	Torch MSE Loss
Optimizer	Adam

Python Torch library was used for model construction, training and testing. The results visualized using Python matplot and seaborn libraries. NVIDIA GPU was be used to increase the speed of training and testing of the model.

3. Data, Preprocessing

In current project, the closing stock prices of Google company in US stock market were used. The stock price data were received through Tiingo cloud server. The raw date were in the form of JSON that were parsed and converted to Panda's data frames. For pre-processing of data, Python Sklearn package was used to map data to a range of [0,1].

3.1. Splitting of Data

After scaling of raw data, the dataset was divided into training set (80%) and testing set (20%). For generation of labeled training set, **sliding window** method with the windows size of 20 data points was used. In this method, each X training (X_training) example has 20 data points and the Y training is the data point immediately after the end of X_training set (point # 21). For generation of the next training example, the windows is slides by one point to generate the next X_training and Y_training examples.

4. Results and discussion

4.1. Comparison of Results with Baseline Study

The results of stock closing price modeling by A. Moghar et al [2] were used as the baseline in the current study. They used RNN – LSTM model to predict the Google stock closing prices. In the baseline paper, only final loss values at the end of training were available and no information on training or testing residual errors were reported. Therefore, the training loss at the end of training period in current study and the baseline study were compared. The final training loss values at the end of 100 epochs in the baseline and both RNN-LSTM and RNN-GRU models in current study were compared and shown in the table 2.

Table2. Comparison of Baseline loss and current study RNN LSTM and GRU models final loss at the end of 100 epochs

GRU and LSTM Parameters	Value
Number of Epochs in all studies	100
Base line final training loss [2]	0.000497
Current study RNN-LSTM final training loss	0.0003136
Current study RNN-GRU final training loss	0.0001236

As it can be seen from the table 2, the value of loss in the models of current study are comparable to the baseline study. Therefore, the model architectures and hyper parameters were kept unchanged for the next phase of study.

4.2. Comparison of Results of RNN-LSTM and RNN-GRU Architectures

Different number of epochs in the range of 100 to 2000 were used to train the RNN LSTM and GRU models. Residuals square mean error (RSME) metric between the actual data and training and testing generated values were used. The calculated training and testing metrics are shown in table 3.

Table 3. Comparison of Baseline loss and current study RNN LSTM and GRU models final loss at the end of 100 epochs

Stock Symbol	Number of Epoch	GRU Training RSME	LSTM Training RSME	GRU Testing RSME	LSTM Testing RSME
Google: Goog	100	21.01	30.71	95.76	224.71
Google: Goog	200	19.25	25.73	75.97	168.64
Google: Goog	300	18.59	22.38	71.45	64.76
Google: Goog	600	18.52	18.82	59.77	48.78
Google: Goog	1000	18.50	18.62	59.40	54.21
Google: Goog	1500	18.46	18.46	42.94	41.78
Google: Goog	2000	17.95	18.45	58.20	44.55
Google: Goog	2500	18.48	16.68	74.55	134.68

Additionally, the plots of RSME values for both training and testing are shown in figure 2. As shown in figure 2, GRU model has lower training error than LSTM method and with increase of epochs after 600, both methods yields very similar RSME. Similarly, the LSTM testing RSME is much higher than GRU testing RSME initially. With increase in epochs numbers, LSTM yields slightly lower RSME in the testing. Afterward, testing RSME in both methods gradually decreases until epoch 1500 in which RSME starts to rise.

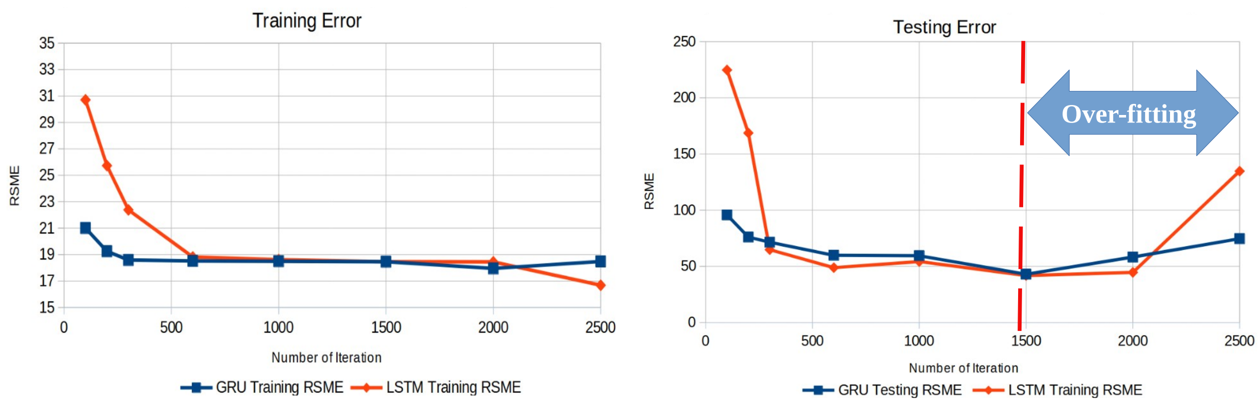

Figure 2. The graphs of residual square errors versus number of iterations in training (left graph) and testing (right graph)

Therefore, we can conclude that the training model starts to over-fit the stock data after epoch 1500. Consequently epoch 1500 was selected as optimum number of epoch for fitting to both LSTM or GRU models on google stock data. The graphs of Google training and actual data as well as training loss graphs are shown in figures 3 and 4.

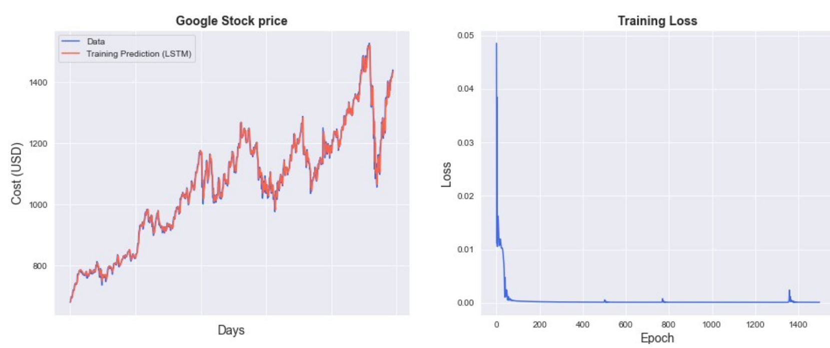

Figure 3. Training and Loss Plot of Google Stock modeled by RNN-LSTM Model

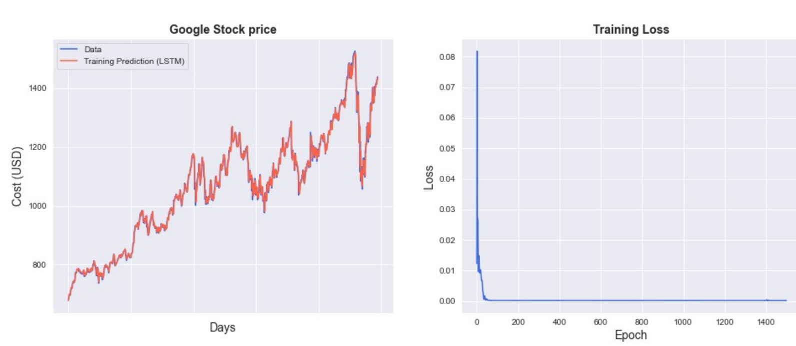

Figure 4. Training and Loss Plot of Google Stock modeled by RNN-GRU Model

After, training of the LSTM and GRU modeles, stock price predictions were done using the trained model. The stock prediction results are shown in figures 5.

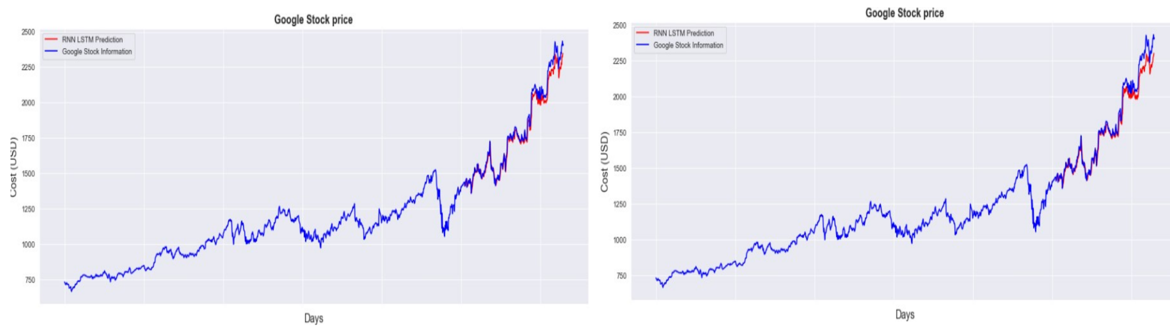

Figure 5. Training and Loss Plot of Google Stock modeled by RNN-LSTM (left) Model and RNN-GRU Model (right)

6. Conclusion

- Two RNN-LSTM and RNN-GRU models were used for training and prediction of Google stock prices. Residual square minimum error (RSME) was used as performance metric in this study.
- GRU model has better performance than LSTM at epoch number lower than 300.
- Both GRU and LSTM models generates comparable RMSE at epoch numbers bigger than 500
- Both GRU and LSTM models demonstrated the sing of over fitting after epoch number 1500. Therefore 1500 was considered for the final training and testing in this study.

7. Suggestion for future work

For the future work, it is suggested to use more advance time series prediction models to predict the stock prices. One of the newest and most advanced time series prediction models is long- and short-term temporal network (LSTNet) that can be used in the model stock prices data which are time series by nature. Another advanced architecture that can be explored are transformers and BERT models that are being used extensively in NLP.

8. References

1. JB Heaton, NG Polson, and Jan Hendrik Witte. Deep learning for finance: deep portfolios, Applied Stochastic Models in Business and Industry, 33(1):3–12, 2017.
2. Adil Moghar, Mhamed Hamiche, Stock Market Prediction Using LSTM Recurrent Neural Network, Procedia Computer Science, Volume 170, 2020, Pages 1168-1173
3. Ryo Akita, Akira Yoshihara, Takashi Matsubara, and Kuniaki Uehara. Deep learning for stock prediction using numerical and textual information. In 2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS), PP1–6. IEEE, 2016.
4. L. Cheng, Y. Huang and M. Wu, "Applied attention-based LSTM neural networks in stock prediction," *2018 IEEE International Conference on Big Data (Big Data)*, 2018, pp. 4716-4718, doi:10.1109/ BigData.2018.8622541.
5. Hao Li, Yanyan Shen, Yanmin Zhu, Stock Price Prediction Using Attention-based Multi-Input LSTM, Proceedings of The 10th Asian Conference on Machine Learning, PMLR 95:454-469, 2018.
6. Chenjie Sang, Massimo Di Pierro, Improving trading technical analysis with TensorFlow Long Short-Term Memory (LSTM) Neural Network, The Journal of Finance and Data Science, Volume 5, Issue 1, March 2019, Pages 1-11
7. Guizhu Shena, Qingping Tana, Haoyu Zhanga, Ping Zenga, Jianjun Xua, Deep Learning with Gated Recurrent Unit Networks for Financial Sequence Predictions, Procedia Computer Science, Volume 131, 2018, Pages 895-903
8. Md.Saiful Islam, EmamHossain, Foreign Exchange Currency Rate Prediction using a GRU-LSTM Hybrid Network, Soft Computing Letters, 26 October 2020, 100009