PREDICTING FOREIGN EXCHANGE USING LSTM AND GRU

May 24, 2020

Gopinath Achuthan
RX44815@umbc.edu
Department of Computer Science and Electrical Engineering
University of Maryland, Baltimore County

Abstract

The goal of this paper is to compare the performance of Long short-term memory model and Gated Recurrent Unit model for Foreign Exchange Rates time series data set obtained from google kaggle: https://www.kaggle.com/brunotly/foreign-exchange-rates-per-dollar-20002019.

Keywords— Long short-term memory, LSTM, Gated Recurrent Unit, GRU, Foreign Exchange, Time Series.

1 Introduction

The motivation of this project is to work with Recurrent Neural Network(RNN) model on time series data set and compare the performance of the different models in RNN. The time series data is the sequence of the events recorded on using the date as an index. The events carry the real number values in its cell. Why we need to analyse the time series data? We analyse the time series data to obtain the meaningful statistical insight from the data. In case of Foreign Exchange Rate, there is a decentralized foreign exchange market to determine it for every country. The aspect of foreign exchange market is to exchange the currency at the current prices. How can we use the Machine Learning model in this time series data? We can use Machine Learning model to do forecasting the foreign exchange rate for various currencies. By forecasting the foreign exchange rate for currencies, we can determine when to exchange the currencies in a way such that we will get the financial gain.

2 Problem Statement

To predict foreign exchange rates with respect to US Dollars. Data from past 20 years for 22 countries will be analysed using LSTM and GRU model of recurrent neural networks. We compare results and performance of two models based on specifications that will be discussed under observation.

3 Approaches and Intuition

To proceed with any machine learning project the important part is data pre-processing. Once done I applied LSTM model and GRU and compared the performance. Let's try to understand intuition behind LSTM and GRU. Simple Recurrent neural networks are derived from feed forward neural networks and use their internal state to process the input. Input is multiplied with previous prediction and passed through tanh prediction. Mathematically simple RNN can be represented as,

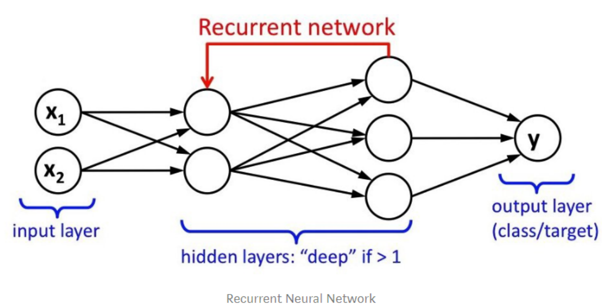


Figure 1: Recurrent Neural Network

3.1 Gated Recurrent Unit

Improved version of standard neural network with a simple tweak we get a gated recurrent neural network from simple RNN, we add 'Update' gate to the existing RNN network. Wherein update operation is nothing but additional mathematical operations with new set of weights. The function of update gate in theory is to determine whether to pass the previous output to the next cell/operation or not.

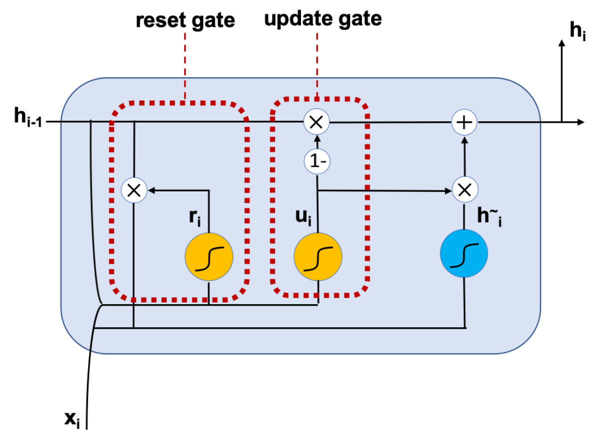


Figure 2: Gated Recurrent Unit(GRU)

3.2 Long short-term memory

We add two more gate in addition to update gate of GRU i.e. forget and output gate. The output of the forget gate tells the cell state which information to forget by multiplying 0 to a position in the matrix. If the output of the forget gate is 1, the information is kept in the cell state. From equation, sigmoid function is applied to the weighted input/observation and previous hidden state. Forget gate decides which out of all the functions which outputs should be moved to the next section

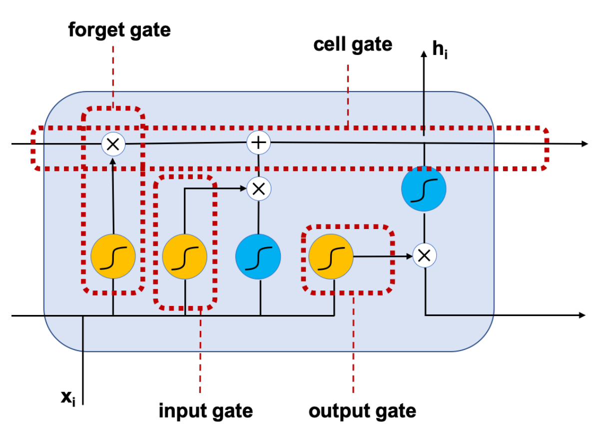
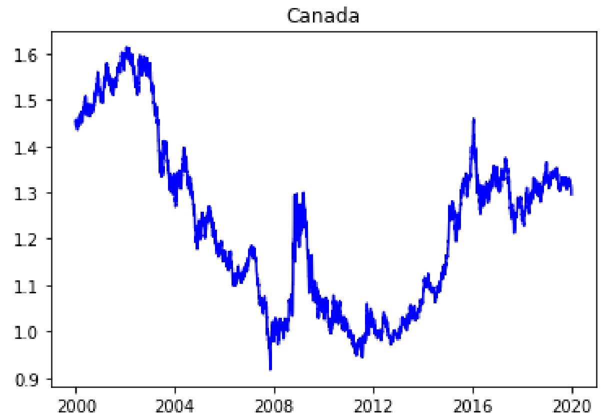
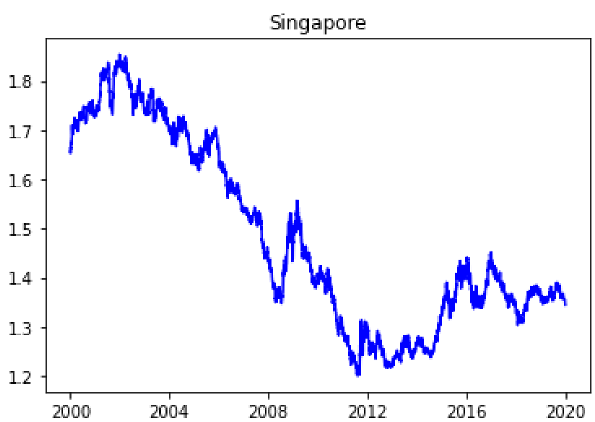


Figure 3: Long short-term memory(LSTM)

4 Experiments

In general, the data scientist work on LSTM and GRU models on time series data set and choose which models works best for there problem. In this project, the time series data set is obtained from google kaggle. In that data set, it carries information about the different countries currency value based on United States currency(dollar) from 2000 to 2019. Based on this information we can forecast the real number currency value. Initially, we need to choose, on which country we are gonna predict the stock exchange value. Now, the data set will be univariant. The stock exchange rate for few countries are given below.





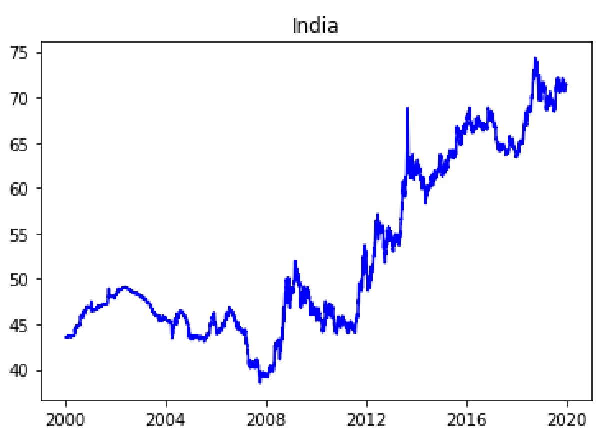


Figure 4: Time Series Data

Let us choose India time series data for this project. As specified above, we need to preprocess the time series data. We can't directly work on data set. Because in this data set, there are few cell is filled as ND, which represent no data on that event. To fill that cell, we use the previous dated data. After filling the data, the next step will be to obtain the data set in such a way the input carries that certain amount information about the previous data. For example, if we are gonna predict the currency value on day 10, the input for the model will be data from day 7 to 9. In this example, we consider the window size as 3. Due to that, we can include the memory details in LSTM and GRU model. In this project, keras library is used for implementing the RNN models. Here, if I mention RNN models, it is referring to the LSTM and GRU model. Now, we observe the performance of the RNN models based on different window sizes. As we know, most of the time, we tend to shuffle the data set and use the machine learning model to do certain statistical or probability operation and perform regression or classification analysis. But here, the data set must not be shuffled, so it maintain the essence of the time series data.

5 Observation

In this project, the task is to forecast or predict the currency rate with respect to dollar. For every machine learning project, we need some metrics to analyse the performance of the model. Here, we are working on the real valued data. It is similar to the regression analysis. Thus, in this project, the most common metric Mean Square Error(MSE) is used to compare the performance of the models. The intuition behind the MSE is to return the real number value representing how far the predicted data is observed from actual data. If the mean squared error close to zero, then the performance of the machine learning is good. The below figure represent the training loss function value of the RNN models for window size as 2 based on the epoch where epoch is one complete presentation of the data set to be learned to a learning machine.

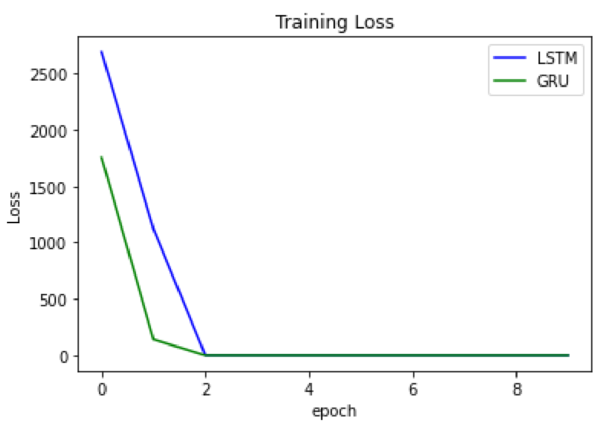


Figure 5: Training Loss in RNN models

For this graph, we can observe that the loss value of GRU model is way lower than LSTM model in first epoch. In third epoch itself, we obtain the MSE value closer to zero in both the model. Thus, few epochs is enough to train both the model for this problem.

Now, let us observe the performance of the RNN models in test loss based on varying number of window size. The observation of the models in form of table and figure are given below.

Window Size	LSTM model	GRU model
1	2.034	1.048
2	0.672	0.052
3	1.36	0.137
4	0.759	0.053
5	0.113	0.124
6	0.092	0.059
7	0.177	0.236

Table 1: Test Loss

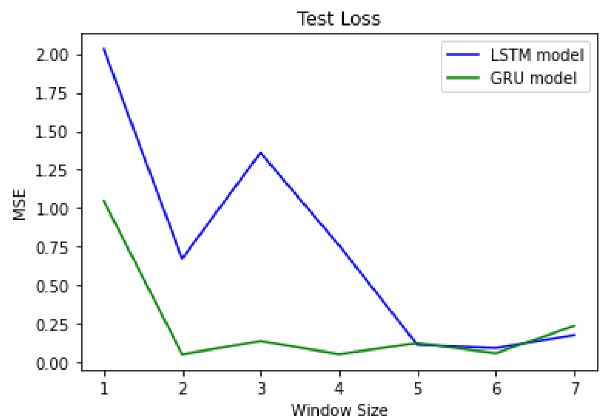


Figure 6: Test Loss in RNN models

From this observation, we can determine the performance of the GRU model is better than LSTM model for this problem. As we increase the value of window size, the loss of both the models are improving. At window size 5, we can see the loss value of LSTM model is lower that GRU model. After that, the performance of both the models is reducing slightly.

Now let us see, the prediction graph of RNN models with window size as 4. From the below figure, we can clearly see how far the prediction value of the model is deviated from the actual value.

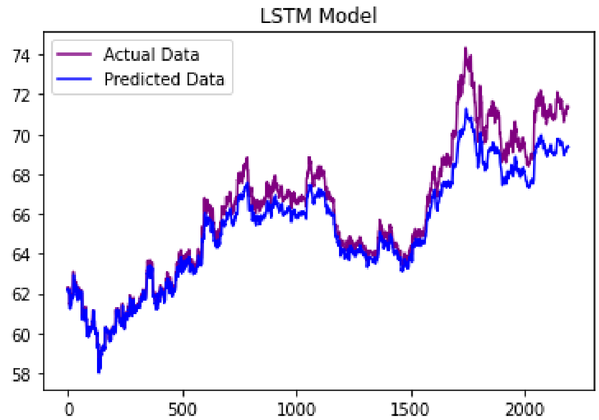


Figure 7: Prediction using LSTM model

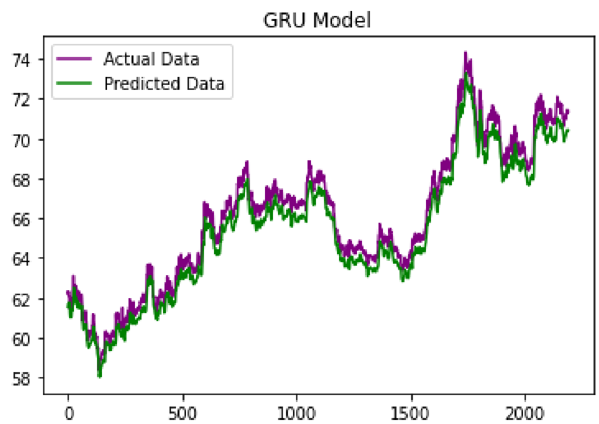


Figure 8: Prediction using GRU model

6 Conclusion

Therefore, from the observation we got from the performance of the LSTM and GRU model, we can say that the performance of the Gated Recurrent Unit model is better than the Long short-term memory model for this problem.

References

[1] Xingjian, S. H. I., et al. "Convolutional LSTM network: A machine learning approach for precipitation nowcasting." Advances in neural information processing systems. 2015.

[2] Gers, Felix A., Nicol N. Schraudolph, and Jürgen Schmidhuber. "Learning precise timing with LSTM recurrent networks." Journal of machine learning research 3.Aug (2002): 115-143.

[3] Nelson, David MQ, Adriano CM Pereira, and Renato A. de Oliveira. "Stock market's price movement prediction with LSTM neural networks." 2017 International joint conference on neural networks (IJCNN). IEEE, 2017.

[4] Merity, Stephen, Nitish Shirish Keskar, and Richard Socher. "Regularizing and optimizing LSTM language models." arXiv preprint arXiv:1708.02182 (2017).

.