Stock Price Prediction Using Machine Learning

Aashay Chaudhari

November 10, 2021

STOCK PRICE PREDICTION USING MACHINE LEARNING

Aashay Chaudhari

Guide: Anjali A. Shejul, Assistant Professor, SCET, MIT, Pune

Abstract

'In the U.S. stock market and many other developed financial markets, about 70-80 percent of overall trading volume is generated through algorithmic trading'
- Fact

Several recent studies have made the use of machine learning in the field of quantitative finance, investment process, as well as, predicting prices of managing and constricting entire portfolio of assets and many other operations that can be covered by machine learning algorithms. In layman's term, machine learning is a term used for all algorithmic methods using computers to reveal patterns based only on the data and without using any programming instructions. For special asset selections and quantitative finance, several machine learning models provide an extensive array of methods that can be used with machine learning to forecast the required future assets value. These type of models offer a mechanism that have the ability to combine weak sources of information and this makes it a strange tool that can be used efficiently.

Neural network studies were originally started in an effort to map the human brain and understand how humans make decisions but algorithm tries to remove human emotions altogether from the trading aspect. What we sometimes fail to

realize is that the human brain is quite possibly the most complex machine in this world and has been known to be quite effective at coming to conclusions in record time.

Introduction

Predicting the future movement of security is the center of the industry of quantitative trading, as the future trading strategy is deployed and created based on our view of the financial market in the future. The trading area has two different methods, namely fundamental analysis, and quantitative trading.
In this project, we have compared some of the existing neural networks (SimpleRnn, GRU, LSTM) and selected the neural network in which we were able to obtain maximum accuracy in quantitative trading.

Objective

1) To predict the closing stock price for the next day using LSTM.

2)To understand the relationship between the number of epochs used to train a model and overfitting.

3)Compare the performance of LSTM against GRU and SimpleRnn

Literature Survey

1) Hybrid Deep learning model (Stock Price Prediction)

Authors: Mohummad Asiful Hussain, Rezaul Karim,Rupa THalushiram, Neil D.B Bruck, Yang Wang ,year-2018(IEEE)

The hybrid model proposed is a combination of LSTM and GRU. The input is passed to the LSTM network first which will generate the first level predictor. The output of the LSTM is later passed to the GRU layer to get a final prediction. This experiment was done on the 500 index historical data of 66 years from 1950 to 2016. Comparing the hybrid models with other models, it outperforms LSTM with 1,2,3,4 layer also. GRU with 1,2 & 3 layers only LSTM with layer 1 is close to a hybrid model.

Model	MSE
LSTM(single layer)	0.001
GRU(single layer)	0.003
LSTM(two-layer)	0.018
GRU(two-layer)	0.001
LSTM(three-layer)	0.003
GRU(three-layer)	0.003
Proposed Model	0.00098

2) Using Neural Networks to Forecast Significant stock price changes

Author: Firoz Kamalov

Date- 9th April 2020

Predicting significant changes,not the actual price or direction of price movement on previous changes. Constructed and tested- MLP(multilayer perceptron), CNN, LSTM.

Relative Strength Index-Used RF and RSI as benchmark

eg.-performed using data on four major are publicly traded companies.

Adjusted daily stock prices from 2009-2019. They observed for 7,14,30,60 day window.

Conclusion: LSTM model yields the best results because it is well situated to analyze sequential data provided. CNN is next best model to performing algorithm. The model's performance improves up to a certain point with an increase in the significance threshold. These models are better at predicting more significant changes as compared to the fewer significant changes.

3) LSTM and Simple RNN comparison in the problem of sequence

Authors:Nouhaila Bensalah, Ayad Habib

year-April 2021

A case study is presented on how different sequence to sequence Deep Learning (DL) models perform in the task of Arabic MT in this paper. It consists of a comprehensive comparison between these models based mainly on: Long Short Term Memory (LSTM), Gated Recurrent Unit (GRU), Bidirectional LSTM (BiLSTM), and Bidirectional GRU (BiGRU) is presented. They have trained the model using sentences of length 20 for both, the source and the output. They have trained the model for 20 epochs with a batch size of 64. Finally, the early stopping technique that is based on the validation loss is added to stop the training process after 4 epochs, if this loss is observed to start increasing. Experimental results show that BiGRU as an Encoder, BiLSTM as a Decoder, and the attention mechanism achieve the best results in terms of BLEU score and computational speed.

4)LSTM vs GRU for Arabic Machine Translation

Authors:Nouhaila Bensalah, Ayad Habib ,year-April 2021

A case study is presented on how different sequence to sequence Deep Learning (DL) models performed in the task of Arabic MT. A comprehensive comparison between these models was done based mainly on: Long Short Term Memory (LSTM), Gated Recurrent Unit (GRU), Bidirectional LSTM (BiLSTM), and Bidirectional GRU (BiGRU) and presented. They trained the model using sentences of length 20 for both, the output and the source. They trained the model for a batch size of 64 with 20 epochs. Finally, the early stopping technique based on the validation loss is added to stop the training process after 4-5 epochs, if this loss is observed to increase. Experimental results showed that BiGRU as an

Encoder, BiLSTM as a Decoder, and the attention mechanism achieved the best results in terms of computational speed and BLEU score.

*Backpropagation

Backpropagation is an algorithm that enables us to update all the weights in the neural network simultaneously. This results in a drastic reduction in the complexity of the process to adjust weights. If we were not using this algorithm, we would be required to adjust each weight individually by figuring out what impact that particular weight has on the error in the prediction.

*Overfitting

The term overfitting is used to refer to a model that models the training data too well.

Overfitting occurs when a model learns the detail and noise in the training data to the extent that it has a negative impact on the performance of the model on new data. This means that the noise or random fluctuations that occur in the training data is picked up and learned as concepts by the model. The problem is that these concepts do not apply to the new data provided and negatively impact the model's ability to generalize.

SimpleRNN:

A Simple Recurrent Neural Network (RNN) is a class of artificial neural networks with connections between nodes from a directed graph along a temporal sequence. This allows the RNN to exhibit temporal dynamic behavior. Being a derivation from the feedforward neural networks, RNNs can use their internal state (memory) to process variable-length sequences of inputs provided. This makes them suitable for tasks such as unsegmented, connected handwriting recognition or speech recognition.

GRU:

Gated recurrent units (GRUs) are a gating mechanism in the recurrent neural networks, introduced by Kyunghyun Cho et al in 2014. The GRU is similar to a long short-term memory (LSTM) with a forget gate, but has fewer parameters than LSTM, as it lacks an output gate. GRU's performance on some tasks such as polyphonic music modeling, speech signal modeling, and natural language processing was found to be similar to that of LSTM.GRUs have been observed to exhibit better performance when it comes to certain smaller and less frequent datasets.

Code:

Model of LSTM (Single Input LSTM):

1) model = Sequential()

2) model.add(LSTM(units=50, return_sequences=True,input_shape=(x_train.shape[1],1)))

3) model.add(LSTM(units=50, return_sequences=False))

4) model.add(Dense(units=25))

5) model.add(Dense(units=1))

Observations

We observe that LSTM continues to increase its closing price prediction accuracy(calculated by RMSE) as we keep increasing the number of epochs over the same dataset from 1. However, the LSTM network starts overfitting after a certain number of epochs(9 in this case) which leads to a decrease in its prediction accuracy. Given below are the number of epochs and dataset size over which the given network is trained along with its corresponding RMSE(error).

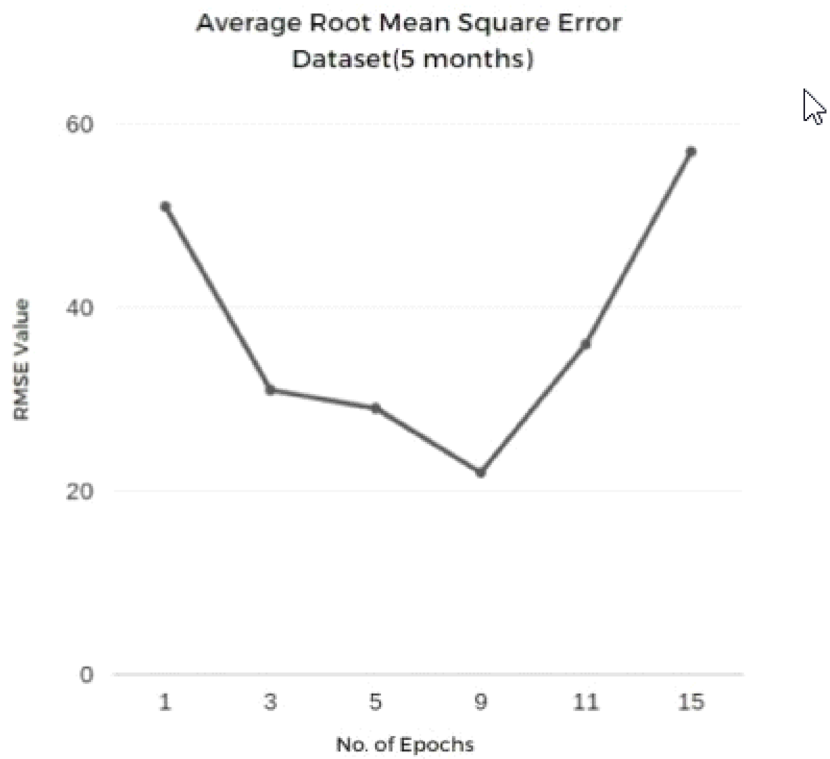
Average Error respective to Epoch variation for LSTM(Single Input LSTM)

No. of years	Epoch		Average Error		% Error
	1	254	323	249	2.39
	3	133	325	150	1.78
10	**5**	215	203	240	1.93
	9	94	122	110	0.95
	11	262	89	93	1.30
	15	221	206	80	1.48

-> The security under consideration for the above observations is the NIFTY50 index. The value of the index was 11438 when these observations were taken.

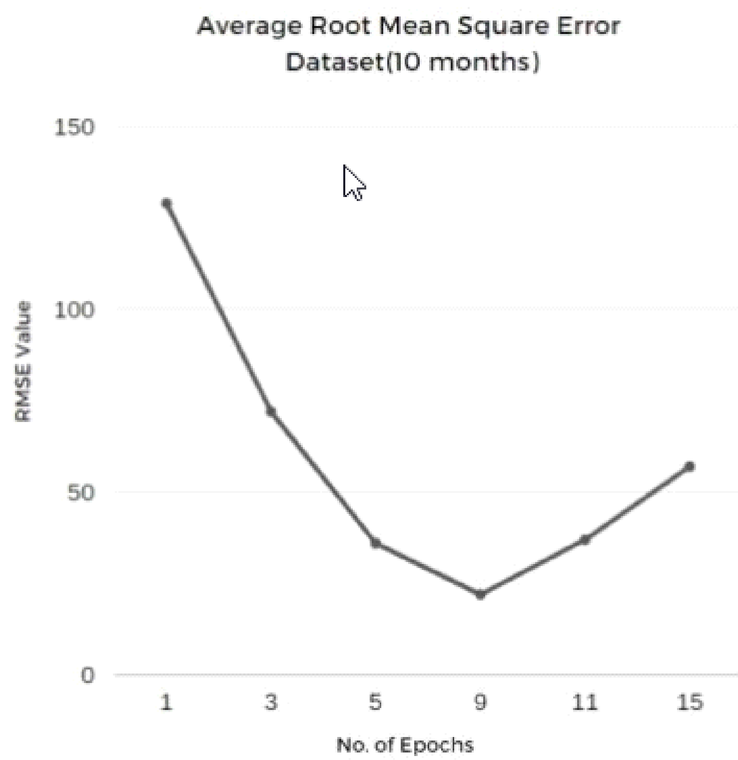
RMSE Error relative to Epoch and dataset time period variation

No. of years	Epoch		RootMeanError		Average
	1	42	70	41	51
	3	30	32	32	31
5	5	28	23	38	29
	9	24	24	18	22
	11	36	40	34	36
	15	60	52	56	57



No. of years	Epoch		RootMeanError		Average
	1	59	71	128	129
	3	85	93	39	72
10	5	65	24	21	36
	9	24	21	21	22

	11	43	42	28	37
	15	66	83	25	57

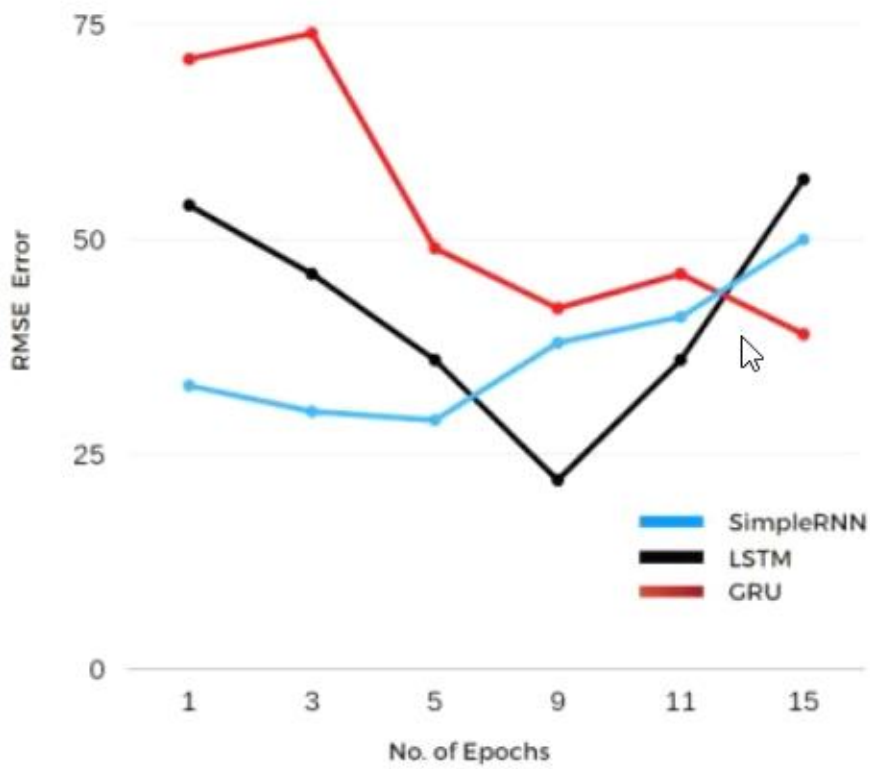


As we can see, the maximum accuracy obtained by our LSTM model defined above is at 9 epochs, provided the learning rate, optimizer, batch size, and all other relevant factors are kept constant. The average error keeps on decreasing till we reach the optimum number of epochs(9 in this case) and then starts increasing again after crossing our optimum number value of epochs. This behavior displays the property of overfitting of a neural network.

Evaluation of LSTM against GRU and SimpleRNN:(Dataset: 5 months)

Epochs	LSTM	RNN	GRU
1	54	71	33
3	46	74	30
5	36	49	29
9	22	42	38
11	36	46	41
15	57	39	50

***comparison made on basis of RMSE values.**



Conclusion

After optimizing our neural networks on various factors, we could reach a minimum average error of 0.9%. This signifies that the neural networks can give you a general idea of tomorrow's stock closing price position based on its training and the data supplied to it. This can be used as another tool in the trader toolkit to give them an idea of how the market might move the next day.

References

[1] Mohummad Asiful Hussain, Rezaul Karim,Rupa THalushiram*, Hybrid Deep learning model used for stock price prediction,* 2016

[2] Firoz Kamalov , *Forecasting Significant stock price changes with the use of Neural Networks*

[3] Yaya Heryadi, Harco Leslie Hendric Spits Warnars, Achmad Imam Kistijantoro *, Comparison between LSTM and Simple RNN in the problem of sequence to sequence on Conversation Data Using Bahasa Indonesia*

[4] Nouhaila Bensalah, Ayad Habib , *LSTM vs GRU for Arabic Machine Translation*