International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-3, January 2020

Indian Stock-Market Prediction using Stacked LSTM AND Multi-Layered Perceptron

Siddharth Banyal, Pushkar Goel, Deepank Grover

Abstract -We aim to construe the Stacked Long–Short term memory (LSTM) and Multi-layered perceptron intended for the NSE-Stock Market prediction. Stock market prediction can be instrumental in determining the future value of a company stock.It is imperative to say that a successful prediction of a stock's future price could yield significant profit which would be beneficial for those who invested in the pipeline of things including stock market prediction. The model uses the information pertaining to the stocks and contemplates the previous model accuracy to innovate the approach used in our paper. The experimental evaluation is based on the historical data set of National Stock Exchange (NSE). The proposed approach aims to provide models like Stacked LSTM and MLP which perform better than its contemporaries which have been achieved to a certain extent. This can be verified by the results embedded in the paper . The future research can be focused on adding more variables to the model by fetching live data from the internet as well as improving model by selecting more critical factors by ensemble learning.

Keywords: Deep Learning,RecurrantNeural Network, Stacked LSTM,Multi layered Perceptron Stock Market Prediction, Mean Squared Error, Mean Absolute Percentage Error.

I. INTRODUCTION

"The future is ours to shape. I feel we are in a race that we need to win. It's a race between the growing power of the technology and the growing wisdomand we need to manage it"

Financial theorists, and data scientists for the better part of the last 50 years, have been employed to make sense of the marketplace in order to increase return on investment. The recent technological advancements in the field of Machine Learning has engendered various Models for Stock prediction.

There always have been an incisive debate on predicting the stock to say the least of course. If we decode the stock market term it allows a person to buy or sell the ownership of a company which we incisively call stocks. If company's profit goes up you earn some profit and if it goes down you incur some losses.

It's quite evident to say that a successful prediction could yield a significant profit. It has led to various research on this topic. So, using Machine learning we can predict the future prices of stocks using the data set of past prices.

Revised Manuscript Received on January 30, 2020.
***** Correspondence Author
Siddharth Banyal*, Maharaja Agrasen Institute of Technology Rohini Sector -22, New Delhi, India.
Pushkar Goel, Maharaja Agrasen Institute of Technology. Rohini Sector -22, New Delhi , India.
Deepank Grover, Maharaja Agrasen Institute of Technology Rohini Sector -22, New Delhi , India.

Our Study aims to (1) To devise a model that predicts the value of the stock from thePrevious data sets. (2) Comparing various other model accuracy with our approach.

The main purpose of the paper is to predict the stock accuracy using the fundamental approach of Stacked LSTM.

In the quest of analyzing and predicting the stock values various approaches has been used. The study by [2] combines Random forest and LSboost for predicting the stock values. They have used technical indicators that provide insights to the expected stock price behavior in future which in turn are used to train random forest. The model in [1] predicts the values of stocks by amalgamating factors like historical stock prices for a particular company , the tweets and the news headline of the country.

The remnant structure of the paper as follows. Section 2 will explain the experimental setup and the rudimentary concepts followed by section 3 which will explain the research methodology and section 4 will include the results of our research. Section 5 will give conclusion of this paper.

2A. Experimental setup

The experiment has been conducted by a decent powered Intel® core ™ i7-7500U CPU @ 2.70GHz (4 CPU's) with a memory size of 8GB. The data-set has been acquired from [8] of Infosys Limited for a period of 52 weeks starting from 4-July-2018 to 7-July-2019. Python has been used as the development language with Development environment being provided by Windows. Anaconda Tools have been used for providing integrated development environment.

II. METHODOLOGY :

MLP:

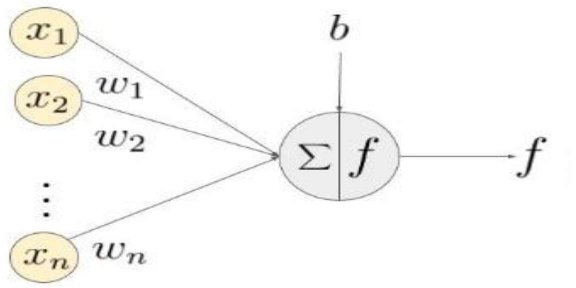
Multilayer perceptron is a case of feed forward ANN(artificial neural network). It has a minimum of 3 layers- a input layer, hidden layer and output layer. There is one input layer, where input is fed, one output layer which gives output and one or more hidden layer. Multiple hidden layers can be incorporated in the model to increase its complexity.

A MLP model consist of various nodes, each of which takes number of inputs $\mathbf{x_i}$ each associated with a weight $\mathbf{w_i}$. The output is given by summation of all products of input and weights which is then added with bias **b.**

The mathematical equation is given as :

$$f = b + \sum_{i=1}^{n} x_i w_i$$

Where x_i : input to the unit
W_i : weight matrix
b : bias to the unit



This model consist of various fully connected layers with each unit of previous layer connected with each unit of next layer.

This leads to each node having a unique set of weights autonomous of other nodes of the same layer.They also may contain several functions chained with each other to perform an affine transformation of linear sum of inputs. There may be problem of overfitting in this model which can be found using loss functions. This can be then handled by using appropriate optimizer functions. This can thus be very useful for predicting data with time stamps with particular trends.

LSTM:

Recurrent Neural networks bring persistence of memory in learning process in contrast to traditional neural networks leading to its profound use in tasks like speech recognition and language model.(references needed). They have a memory cell which is used to store historical information , instrumental in finding trends and relations in the data.

However, there is problem of vanishing gradient in RNN where there is exponential decay of loss function with time. This leads to difficulty in training models with long term temporal dependency because information in 'memory' is stored over time in the model. LSTM overcomes this problem by using special units in addition to memory units comprised of multiple gates. These gates control when information enters, when it outputs and when information is forgotten. This allows the model to learn long term dependencies as these gates overcome vanishing gradient problems.

A LSTM model takes three inputs, current input of time step i_t, output of previous time step O_{t-1} and memory of previous time unit C_{t-1}. The output of LSTM model is the memory unit C_t and the output O_t. There are various hidden vectors use in the working of LSTM H_t. The variables are transformed under sigmoid and hyperbolic tangent functions σ_g with weight and bias matrix W. various operations undertaken in the model is as follows :

O_t : *Output gate activation vector*

$$f = \sigma_g(w_f x_t + U_f C_{t-1} + B_f)$$

$$i_t = \sigma_g(w_i x_t + U_i C_{t-1} + B_i)$$

$$O_t = \sigma_g(w_o x_t + U_o C_{t-1} + B_o)$$

$$C_t = f_t \, o \, C_{t-1} + i_t \, o \, \sigma_c(w_c x_t + B_c)$$

$$h_t = \sigma_h(O_t \, o \, C_t)$$

x_t : *input vector to the LSTM*

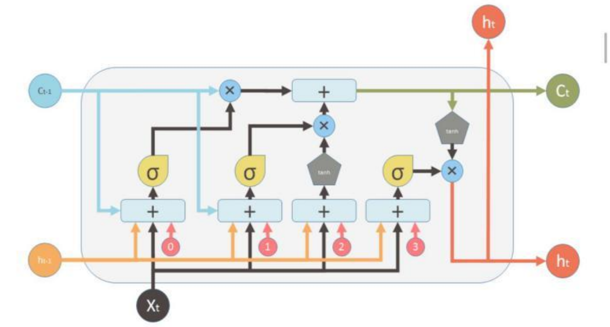
f_t : *forget gate activation vector*

i_t : *input gate activation vector*

h_t : *hidden state vector* c_t : *cell state vector*

W: *weight matrices and bias parameters*

σ : *sigmoid and hyperbolic tangent function*



LSTM model can be considered as a pipe with two valves- forget valve which decides whether the information from previous steps is to be retained or not and the memory valve which stores information of previous steps and is in integrated with forget valve to take its output into consideration as well. This model is thus able to work properly on data with long term temporal dependency as well because forget gate can remove irrelevant features and thus prevent exponential decay.

III. PROPOSED MODEL

Data Prepossessing

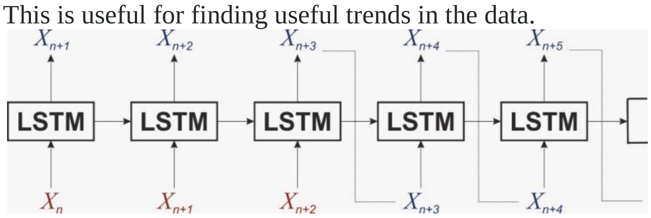
This model employs normalization technique to scale the features in range of zero to one . Normalization converges the data into a common scale without distorting the relative difference of values. This is done to ensure that mathematical magnitude of a feature does not affect the weight age of feature in the model.

$(x_i$ - mean(x)/stdev(x))

Time Stamping

Time stamping is done to predict the value X_{T+1}value in a data set from [X_0,X_1,X_2,. . . . . ,X_T] ,in the context of a date-stream in real-time.Considering the data for 60 observation , we find the predicted the data for the 61th observation .

This is useful for finding useful trends in the data.



Activation Function

It is used to get the output of node and is determine the output of neural network and a maps the value between 0-1. It is also known as Transfer Function.The activation function used in my model is rectified linear unit because of the non- linear nature of stock market data. It is also very efficient when stacked layers are considered.

R(x) = max(0,Z)

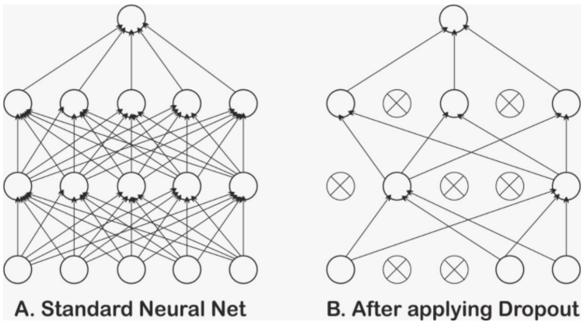
Adam Optimizer.

Adam optimizer is used because it takes advantage of momentum to calculate gradient instead of traditional methods of finding gradient like Stochastic Gradient descent. It uses adaptive learning to find the most appropriate result sooner than the conventional methods.

Dropout layer.

In this approach we simply remove some of the neurons in the neural network randomly so as to lessen the possibility of over fitting and reduce the computational cost of the model. It helps in garnering

attention around more robust features and lessening the weight age of typical features.



We have used keras library to implement our model. Keras is a framework build upon Tensor flow which is very helpful in writing efficient and user friendly code .

IV. RESULTS:

Sequential API is used in keras for building the LSTM model.

Mean square error is deployed in the results to find the deviation between the estimated value and what is actually predicted.A straight line that passes through all the points which fits them in the best way. This line contains the predicted points. The general formula is given below .

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \tilde{y}_i)^2$$

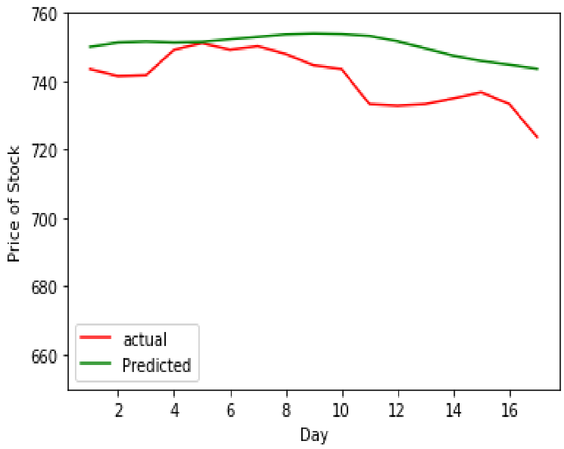
We have used dense layer to accumulate the result of previous layers into a single probability. This leads to easier computation cost and finding accurate results.

The results of the cognitive techniques are as follows.

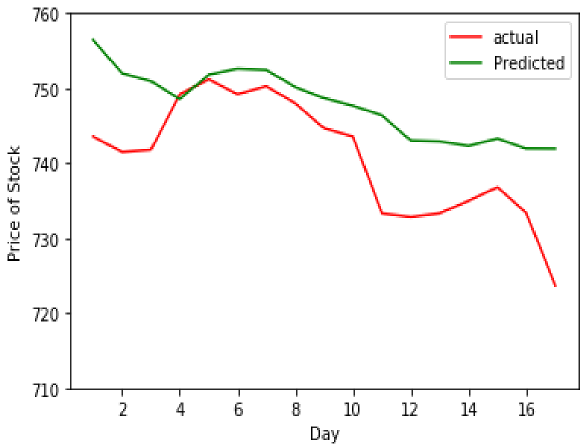
	LSTM				MLP			
Timestamp	Adam		RMSprop		Adam		RMSprop	
	Sigmoid	Relu	Sigmoid	Relu	Sigmoid	Relu	Sigmoid	Relu
30	2.82%	0.73%	2.96%	0.67%	3.93%	0.64%	8.01%	20.67%
45	0.85%	1%	1.45%	0.80%	1.41%	0.95%	0.86%	4.49%
60	1.05%	0.98%	2.32%	0.58%	0.94%	0.70%	20.59%	20.67%

The above table illustrates the percentage difference of models - LSTM and MLP under diametric optimizer , Activation function and Time-stamps.It can be inferred that leading accuracy is given by Stacked LSTM with 60 Time-stamp, RMS Optimizer and RElu

The above table illustrates the percentage difference of models - LSTM and MLP under diametric optimizer , Activation function and Time-stamps. It can be inferred that leading accuracy is given by Stacked LSTM with 60 Time-stamp, RMS Optimizer and RElu Activation function.



Actual vs Predicted value of stock using LSTM



Actual Vs Predicted value of stock using MLP

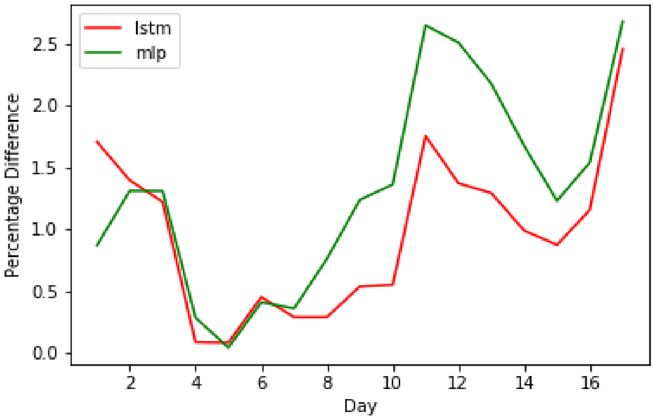


Fig : percentage differences amongst LSTM and MLP.

The above graph provide conclusive evidence of the accuracy of the incisive approach explained in the paper. The first two graph indicate the difference in the predicted and the actual values in an absolute scale. The percentage difference of the actual and predicted values lie in the range of 0-5 %.

The third graph shows the percentage difference of the predicted and actual values of the approach of LSTM and MLP on a relative scale. It illustrates that both models perform better than its consociate model in particular areas but it can be observed that LSTM outperforms MLP in most of the regions. This thus opens further avenue for research in ensemble model picking on beneficial features of both the approaches.

V. CONCLUSION

This paper, proposed a machine learning model that uses stacked LSTM and MLP to compare their accuracy. The model helps us predicting the value of stock market given the non-linear nature of the stock market which is complex in nature. Stock market brokers can use this approach to predict the value of a stock in an efficient manner.

This model can predict values of stock market using previous trends and data .It can be further improved by incorporating external factors which are separate from mathematical figures but affect the prices nevertheless. Further we can make ensemble model to acquire the favourable traits of various model. For further improvement we can integrate it with a dynamic web scraping API which can collect information about changes happening in the real time and fully automate this process.

REFERENCES

1. M.Hegde, G. Krishna and R. Srinath, "An Ensemble Stock Predictor and RecommenderSystem", *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 2018.
2. B. Narayanan and M. Govindarajan, "Prediction of Stock Market using Ensemble Model", *International Journal of Computer Applications*, vol. 128, no. 1, pp. 18-21, 2015.
3. Yaya Heryadi ; Harco Leslie Hendric Spits Warnars"Learning temporal representation of transaction amount for fraudulent transaction recognition using CNN, Stacked LSTM and CNN-LST." 2017 IEEE INTERNATIONAL Conference on Cybernetics and Computational Intelligence.
4. Harikrishnan, N. and Soman, K. (2018). Detecting Ransomware using GURLS. *2018 Second International Conference on Advances in Electronics, Computers and Communications (ICAECC)*.
5. "Stock Market Prediction Using Hidden Markov Model"Published in: 2014 IEEE 7th Joint International Information Technology and Artificial Intelligence Conference
6. Sharma, A., Bhuriya, D. and Singh, U. (2017). "Survey of stock market prediction using machine learning approach" 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA).
7. Fonseka, C. and Liyanage, L. (2008). A Data mining algorithm to analyse stock market data using lagged correlation. *2008 4th International Conference on Information and Automation for Sustainability*.
8. Nseindia.com. (2019). *NSE - National Stock Exchange of India Ltd.*. [online] Available at: http://www.nseindia.com/ [Accessed 18 Oct. 2019].
9. Towards Data Science. (2019). *Towards Data Science*. [online] Available at: https://towardsdatascience.com/machine-learning-techniques-applied-to-stock-price-prediction-6c1994da8001.[Accessed 5 Sept. 2019].
10. Towards Data Science. (2019). *Towards Data Science*. [online] Available at:https://towardsdatascience.com/machine-learning-techniques-applied-to-stock-price-prediction-6c1994da8001. [Accessed 14 Oct. 2019].
11. Brownlee, J. (2019). *When to Use MLP, CNN, and RNN Neural Networks*. [online] Machine Learning Mastery. Availableat:https://machinelearningmastery.com/when-to-use-mlp-cnn-and-rnn-neural-networks/ [Accessed 1 Oct. 2019].
12. Rao, A. (2019). *Convolutional Neural Network (CNN) Tutorial In Python Using TensorFlow |Edureka*. [online] Edureka. Availableat:https://www.edureka.co/blog/convolutional-neural-network/[Accessed 4 Oct. 2019].
13. Learning, D. and Memory, E. (2019). *Essentials of Deep Learning : Introduction to Long Short Term Memory*. [online] AnalyticsVidhya.Availableat:https://www.analyticsvidhya.com/blog/2017/12/fundamentals-of-deep-learning-introduction-to-lstm/ [Accessed 3 Oct. 2019].
14. Medium. (2019). *Simple RNN vs GRU vs LSTM :- Difference lies in More Flexible control*. [online] Available at: https://medium.com/@saurabh.rathor092/simple-rnn-vs-gru-vs-lstm-difference-lies-in-more-flexible-control-5f33e07b1e57 [Accessed 15 Sept. 2019].
15. Colah.github.io. (2019). *Understanding LSTM Networks -- colah's blog*. [online] Available at: https://colah.github.io/posts/2015-08-Understanding-LSTMs/ [Accessed 18 Oct. 2019].
16. Medium. (2019). *Activation Functions in Neural Networks*. [online] Available at: https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6 [Accessed 11 Sept. 2019].
17. Human Object Identification for Human-Robot Interaction by Using Fast R-CNN" 2018 Second IEEE International Conference on Robotic Computing (IRC)".
18. LSTM Neural Networks for Baybáyin Handwriting Recognition" 2019 IEEE 4th International Conference on Computer and Communication Systems (ICCCS)"
19. A Runge-Kutta MLP Neural Network Based Control Method for Nonlinear MIMO Systems
20. Using Support Vector Machines to Classify Student Attentiveness for the Development ofPersonalizedLearningSystems" 2013 12th International Conference on Machine Learning and Applications"

AUTHORS PROFILE

Siddharth Banyal, Siddharth Banyal is a B. tech undergraduate pursing majors in Information Technology from Maharaja Agrasen Institute of Technology,IPU,Delhi. Having worked on various cognitive Machine learning models such as image classification/detection/localization/analysis and handwriting recognition. Experienced in web Development and has provided solutions to various local business. Currently engrossed in learning Android Development so as to envisagehis machine learning models in mobile applications.

Pushkar Goel , Pushkar Goel is a Student, pursuing B.tech with major in Computer Science from Maharaja AgrasenInstitute of technology,IPU,Delhi, having interest in Machine learning,maths and real-life problemsolving. He Love to program and learn new technologies and have Experience in webdevelopment, machine learning model deployment using Python and various data structures like linked list,

stacks and graph. Goel has worked in various fields of machine learning like Naturallanguage processing, image processing,time series data analysis and real-life categorization problem. His current research focuses on Stock market analysis and Prediction using deep learning techniqueslike Long short term memory(LSTM) and Multilayer Perceptron.

Deepank Grover , A student pursuing B.Tech major in information technology from maharaja Agrasen Institute of Technology, IPU, Delhi. He is interested in doing real life and challenging projects with the scope of python,machine learning and deep learning. His research interests include processing time series data, Stock evaluation and to work on real time information. He is responsible and ambitious student with a