iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 Issue: VI Month of publication: June 2020

DOI: http://doi.org/10.22214/ijraset.2020.6293

Stock Price Trend Prediction

Sahil Bhatia[1], Srishti Kishore[2], Siddharth Nager[3], Mr. Anurag Agarwal[4]
[1, 2, 3]Students, [4]Assistant Professor, Department of Information Technology, Bharati Vidyapeeth's College Of Engineering New Delhi, India

Abstract: Stock worth measures have reliably pulled in the thought of various agents and authorities. Standard speculation holds that stock trades are sporadic in nature, and it is senseless to endeavour to envision them. Since various components are incorporated, anticipating the stock worth itself is a troublesome issue. Until further notice, the market continues like a popularity based machine, yet as time goes on, it acts like a measuring machine, so it can foresee exhibit designs over a progressively broadened time period. The utilization of machine learning advancement and various counts in stock worth examination and figure is a promising field. In this paper, we first rapidly plot the logical classification of the stock markets and taxonomy of stock market prediction methods. By then, we will focus on some investigation achieves stock assessment and deciding. We look at the particular, technical, fundamental, and long short term memory network strategies used in stock examination. Finally, we present a couple of challenges and investigation opportunities in this field.

Keywords: Machine Learning algorithms, Supervised Learning, Unsupervised Learning, Supervised Learning algorithms, Support Vector Machine (SVM), Support Vector Regression (SVR), Linear Regression, Logistic Regression, Long Short Term Memory network (LSTM).

I. INTRODUCTION

Financial market is one of the most fascinating inventions of our time. They significantly influence various fields, for instance, business, preparing, work, development, and thusly significantly influence the economy. Consistently, monetary authorities and researchers have been enthused about making and testing stock worth direct models. In any case, considering the way that the market is dynamic, non-straight, non-fixed, non-parametric, clamorous and disordered, it is incredibly trying to dismember the eccentrics and worth direct of the protections trade. The protections trade is impacted by various significantly relevant parts, including fiscal, political, mental, and companions express factors. Particular and key examination are the two essential techniques for analyzing financial markets. To place assets into stocks and get high advantages with OK, theorists have used these two essential methods to choose decisions in budgetary markets. Monetary trade esteem desire is a questionable business. Consistently, a couple of speculations about the protections trade have been conceptualized. They either endeavour to explain the possibility of the protections trade or endeavour to explain whether they can beat the money related trade. We discussed potential troubles and possible future assessment headings. Section 2 gives various approaches towards stock market analysis. Section 3 presents the composing research and studies on the stock market, and Segment 4 discussions about research gaps in the procedures referenced in Section 3. Section 5 outlines way to deal with address troubles and maybe beat the market by LSTM approach and study its strong structure. Results and implementation are discussed in Section 6. Finally, Section 7 summarizes the paper by a conclusion.

II. APPROACHES TO STOCK MARKET ANALYSIS

The most recent improvements in stock investigation can be separated into three classifications: measurements, design acknowledgment and machine learning (ML). The greater part of these classes have a place with the more extensive specialized examination classification, yet some machine learning advances additionally consolidate the more extensive specialized investigation class with fundamental investigation strategies to foresee the securities exchange. These procedures have picked up fame in the ongoing field of stock examination and have demonstrated empowering results. Hence, we will talk about the machine learning procedures recently utilized in it. Machine learning has widely considered its potential in anticipating money related markets. Machine learning errands are generally partitioned into administered learning and solo learning. In administered learning, a lot of named input information and watched yield information can be utilized to prepare the calculation. In any case, in unaided learning, just unlabeled or watched yield information is accessible. The objective of administered learning is to prepare a calculation to naturally outline information to given yield information. In the wake of preparing, the machine will figure out how to see the info information focuses and anticipate the normal yield. The objective of solo learning is to prepare calculations to discover examples, connections or groups in a given informational collection. It can likewise fill in as a forerunner to administered learning assignments. A few calculations are utilized in stock value heading forecast. Simple methods such as the single decision tree, discriminant analysis, and Naive Bayes have been substituted by better-performing algorithms such as Random Forest, logistic regression, and neural networks. With nonlinear, data-driven, and easy-to-generalize features, multivariate analysis over the usage of deep Artificial Neural Networks (ANN) and Long Short Term Memory (LSTN) has developed a main and prevalent examination tool in the monetary market analysis. Newly, deep nonlinear neural network topology is opening to fascinate attention in time series prediction.

III. RELATED WORK

Numerous machine learning methodology have been examined for stock worth course desire Ballings et al. [3]. ANN and Support Vector Regression (SVR) are two comprehensively used machine learning figuring for envisioning stock expense and budgetary trade record regards Patel et al. [8]. Controlled learning methods like Support Vector Machine (SVM) and Decision Trees can make sense of how to envision protections trade expenses and examples reliant on recorded data and give noteworthy assessment of bona fide cost. Bernal et al. [1] executed a subclass of Recurrent Neural Network (RNN) known as Echo State Networks (ESN) to anticipate S&P 500 stock expenses using cost, moving midpoints, and volume as features. The methodology defeats the Kalman Filter technique with a little test slip-up of 0.0027. To summarize and support their result, Bernal et al[1] reviewed the estimation on 50 unique stocks and uncovered that their results performed well against front line strategies. Ballings et al. [3] benchmark outfit techniques including Random Forest, AdaBoost, and Kernel Factory against single classifier models, for instance, Neural Networks, Logistic Regression, Support Vector Machines, and K-Nearest Neighbour using data from 5767 uninhibitedly recorded European associations. The makers used on different occasions two-cover cross-endorsement and Zone Area Under the Curve (AUC) as an introduction measure for foresee long stretch stock worth bearing and uncovered Random Forest as the top computation. Milosevic[6] proposed an approach for long stretch gauge of money related trade costs through a request task where a stock is 'adequate' if the stock cost increases by 10% in a year else it is a 'horrendous'. In addition, Milosevic[6] played out a manual part decision, picked 11 noteworthy essential extents, and applied a couple of machine learning computations to stock estimate. It follows that Random Forest achieved the best F-Score of 0.751 against techniques, for instance, SVM and Naive Bayes. Another technique that has Long Short Term Memory (LSTM) orchestrate have shown a lot of assurances for time course of action conjecture Di Persio and Honchar [5] applied three different Recurrent Neural Network models specifically a fundamental RNN, the LSTM, and the Gated Recurrent Unit (GRU) on Google stock expense to survey which variety of RNN performs better. It was evident from the results that the LSTM beat various varieties with a 72% precision on a five-day horizon and the makers moreover explained and demonstrated the covered components of RNN. Roondiwala et al.[8] completed a LSTM framework to foresee clever expenses with features like OHLC. Their results show that the LSTM achieves a RMSE of 0.00859 for the test data to the extent step by step rate changes. Yang et al. [9] proposed an outfit of multi-layer feed forward frameworks for Chinese stock figure. Three portion frameworks were readied using getting ready counts like back propagation and Adam. The troupe was encircled using the pressing methodology Efron and Tibshirani. The results procured display that the Chinese markets are generally obvious and achieve an acceptable precision, exactness, and audit. Zhang et al. [10] propose a stock worth example desire system that can anticipate both stock worth turn of events and its time span (or abatement) rate inside predefined conjecture lengths. They arranged a sporadic woodlands model from bona fide data from the Shenzhen Development Endeavour Market (China) to amass various fastens of stocks into four essential classes (up, down, level, and dark) as demonstrated by the conditions of their close by costs. Their appraisal shows that the proposed system is energetic to the market unsteadiness and outmanoeuvres some current conjectures methods with respect to precision and return per trade. Hossain et al[11] propose a significant learning-based cross variety model that includes two striking DNN structures: LSTM and GRU. The makers arranged a desire model using S&P 500 time course of action dataset spreading over around 66 years (1950 to 2016). The procedure incorporates passing the data to the LSTM framework to create a first level conjecture and a short time later passing the yield of LSTM layer to the GRU layer to get the last desire. The proposed mastermind achieved a Mean Squared Error (MSE) of 0.00098 in desire with beating past neural framework moves close. Starting late, Lv et al. falsely surveyed distinctive ML computations and viewed the step by step trading execution of stocks under trade cost and no trade cost. They utilized 424 S&P 500 record part stocks (SPICS) and 185 CSI 300 Index Component Stocks (CSICS) some place in the scope of 2010 and 2017 and differentiated traditional machine learning computations and impelled significant neural framework (DNN) models. The customary machine learning counts are SVM, Random Forest, Logistic Regression, naïve Bayes, Classification and Regression Tree (CART), and eXtreme Gradient Boosting while the DNN architectures include Multilayer Perceptron (MLP), Deep Belief Network (DBN), Stacked Autoencoders (SAE), RNN, LSTM, and GRU. Their results show that standard machine learning estimations have a prevalent presentation in a huge part of the directional appraisal pointers without considering the trade cost, in any case, DNN models show better execution considering trade cost.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue VI June 2020- Available at www.ijraset.com

IV. RESEARCH GAPS

Machine learning procedures like Support Vector Regression (SVR) ,Linear Regression(LR) and Random Forest(RF) are used for stock price prediction. Artificial Neural Networks (ANN) and Recurrent Neural Networks(RNN) are two Deep learning techniques which are used for predicting price and stock market index values. These models have great results but lags where in forecasting precision and score.

Existing systems report attempts various things with high estimate scores by picking fitting intervals of time for preliminaries to secure elevated standard scores. Exactly when the working condition changes, the current structure doesn't function admirably. The conjecture precision of the current system can be improved. We see that the energy research waits behind the correct execution of the LSTM model, which can be promotion libbed and can be used to all the more promptly anticipate time-sharing issues, for instance, stock worth desire.

V. LSTM ARCHITECTURE

A. An overview of Recurrent Neural Network (RNN)

In a classical neural network, the last yield on occasion fills in as the accompanying yield, anyway if we revolve around authentic wonders, we will find that all around, the keep going yield relies upon the external data, yet what's more on the earlier yield. For example, when people read a book, the cognizance of each sentence relies upon the current word list, yet what's more on the perception of past sentences or the setting made using past sentences. Individuals don't think without any planning reliably. When examining this article, you will see each word subject to your cognizance of the past words. Ordinary neural frameworks can't use the thoughts of "setting" or "persistency". The inability to use setting based reasoning has become a critical limitation of customary neural frameworks. Recurrent Neural Network (RNNs) facilitates this obstruction insightfully. RNNs are associated with inside analysis circles to allow productivity of information. Figure 1 shows a direct RNN with an info circle and its all-inclusive equivalent adjustment one close to the next.

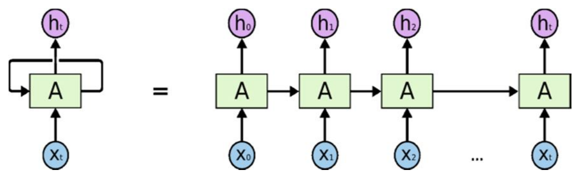

Figure 5.1: An unrolled recurrent neural network [19]

From the start (at time step t), RNN produces ht yield for some information Xt. In at whatever point step (t + 1), the RNN takes two information sources Xt + 1 and ht to deliver the yield ht + 1. Circumnavigating grants information to go beginning with one phase of the framework then onto the following. RNN isn't without its obstacles. When the "particular condition" is close the past, it is significant for right yield. In any case, when RNN must rely upon far away "setting" (that is, something taken in a long time back) to convey the correct yield, it will bomb pitiably. Hochreiter [8] and Bengio and others inspected this limitation of RNN in detail. [9]. They similarly followed back to key points to fathom why RNN may not work eventually. Luckily LSTM hopes to vanquish the recently referenced issues.

B. LSTM Networks

Long Short-Term memory is one of the best RNNs architecture. LSTM presents a limit unit, which is an enlisting unit that can supersede ordinary phony neurons in the covered layer of the framework. Using these limit units, the framework can satisfactorily relate the memory and make remote commitment to time, so it is fitting for logically understanding the structure of data after some time with high judicious limit. A while later, various experts improved the leading work in [11] [12] [13] [14]. After some time, LSTM has been refined to alleviate the drawn out dependence issue. [15] [16] explains the headway and improvement of RNN from LSTM. An irregular neural framework is a kind of reiterating module chain of a neural framework. In a standard RNN, this repeating module has an essential structure, for instance, a single tanh layer, as showed up in Figure 2.

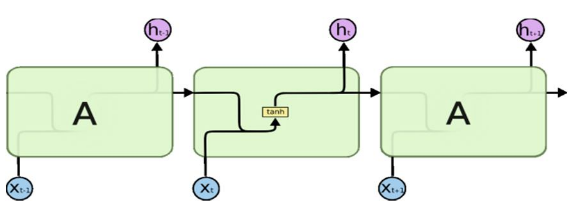

Figure 5.2: The repeating module in a standard RNN contains a single layer [19].

LSTM follows this chain structure; anyway the repeating module has a substitute structure. Rather than having only a solitary neural framework layer, there are four layers that impart in an extraordinarily uncommon way, as showed up in Figure 3.

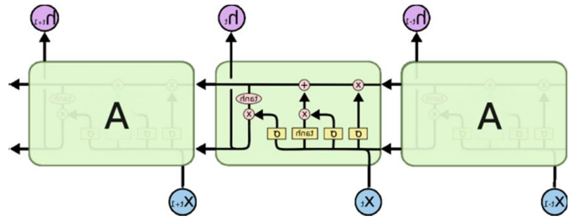

Figure 5.3: The repeating module in an LSTM contains four interacting layers [19].

In Figure 3, each line addresses the entire component vector from the yield of one centre point to the commitment of another center. Pink circles address point-by-point undertakings, for instance, vector extension, while yellow boxes address the academic neural framework layer. The combined lines show connect, while the forked lines exhibit that their substance are being imitated, and the copies are in different regions.

VI. RESULT AND DISCUSSION

The historical data of 6 years of GOOGL stock was collected from finance.yahoo.com and this historical data is used for the prediction of future stock prices trend. The dataset was divided into two parts for training and testing purposes.
The projected LSTM based model is executed via Python using training dataset. In Table3.1 the Losses occurred in executing model with diverse numbers of Epochs are shown:

TABLE 6.1: Number of Epochs vs Loss

Number of Epochs	Loss
50	0.0024
80	0.0013
100	0.0014
120	0.0011

A. Visualization and Comparison of Results

Below are the predicted trends of stock prices for 20 days in comparison with their real price.

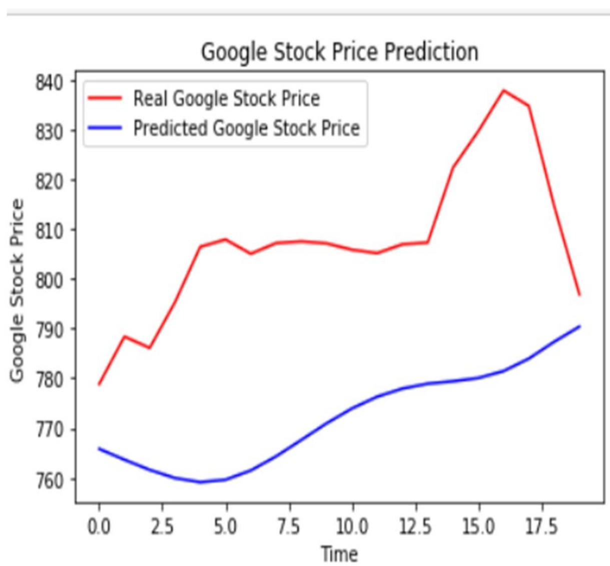

Figure 6.1: Graph for 50 Epochs model

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue VI June 2020- Available at www.ijraset.com

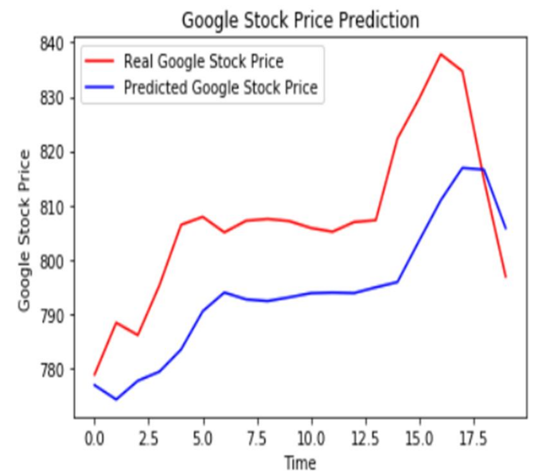

Figure 6.2: Graph of 80 Epochs model

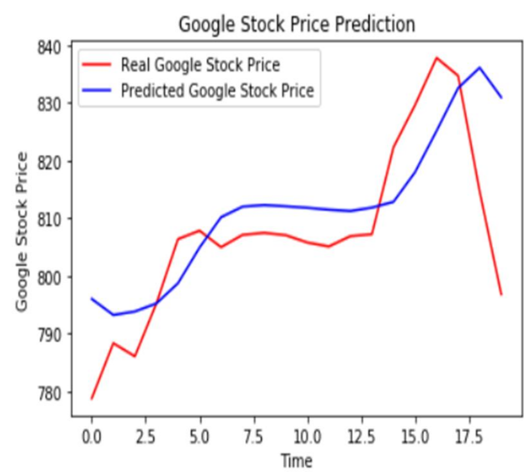

Figure 6.3: Graph of 100 Epochs model

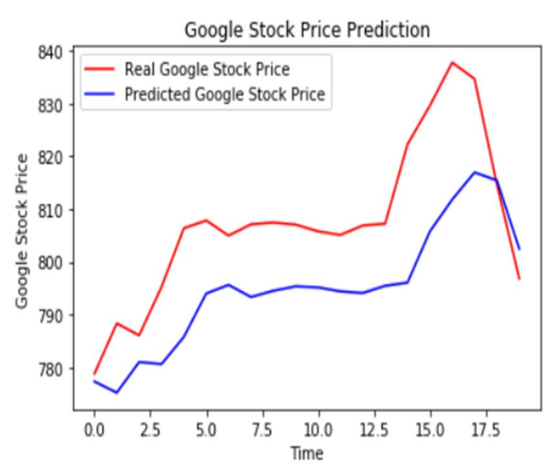

Figure 6.4: Graph of 120 Epochs model

By analysis of above graph we get and the prediction graph of model with 100 Epochs shows the best results i.e. the prediction trend line is most closest to the real trend line .Henceforth it can provide the investor a correct and precise idea nearby the future of a stock by analysing its previous data.

VII. CONCLUSIONS

The acclaim of stock market trade trading is growing, which urges researchers to use new headways for predicting through new techniques. These techniques can bolster experts and any person who deals with the stock trade. To help anticipate stock records, a forecasting model with incredible exactness is required. In this work, we have used the most exact assessing methodologies, to be explicit the use of Recurrent neural network and Long Short Term Memory units, which can give investors, specialists, analysts or anyone motivated by the stock trade with information about future conditions.

REFERENCES

[1] Bernal,Armando, SamFok andRohit Pidaparthi. 2012. Financial Market Time Series Prediction with Recurrent Neural Networks. State College: Citeseer.
[2] Babu, M. Suresh, N. Geethanjali, and B. Satyanarayana. 2012. Clustering Approach to Stock Market Prediction. International Journal of Advanced Networking and Applications 3: 1281.
[3] Ballings, Michel, Dirk Van den Poel, Nathalie Hespeels, and Ruben Gryp. 2015. Evaluating multiple classifiers for stock price direction prediction. Expert Systems with Applications 42: 7046–56.
[4] Dey, Shubharthi, YashKumar, Snehanshu Saha, and Suryoday Basak. 2016. Forecasting to Classification: Predicting the Direction of Stock Market Price Using Xtreme Gradient Boosting. Working Paper.
[5] Di Persio, Luca, and Oleksandr Honchar. 2017. Recurrent Neural Networks Approach to the Financial Forecast of Google Assets. International Journal of Mathematics and Computers in simulation 11:7–13.
[6] Milosevic, Nikola, Equity Forecast: Predicting Long Term Stock Price Movement Using Machine Learning.
[7] Mittal, Anshul, and Arpit Goel. 2012. Stock Prediction Using Twitter Sentiment Analysis.Standford University, CS229. Available online: http://cs229.stanford.edu/proj2011/GoelMittalStockMarketPredictionUsingTwitterSentimentAnalysis.pdf
[8] Roondiwala, Murtaza, Harshal Patel, and Shraddha Varma. 2017. Predicting Stock Prices Using Lstm. International Journal of Science and Research (IJSR) 6: 1754–56.
[9] Yang, Bing, Zi-Jia Gong, and Wenqi Yang. 2017. Stock Market Index Prediction Using Deep Neural Network Ensemble. Paper Presented at the 2017 36th Chinese Control Conference (CCC), Dalian, China, July 26–28.
[10] Zhang, Jing, Shicheng Cui, Yan Xu, Qianmu Li, and Tao Li. 2018. A novel data-driven stock price trend prediction system. Expert Systems with Applications 97: 60–69.
[11] Y. Bengio, P. Simard, P. Frasconi and others, "Learning long-term dependencies with gradient descent is difficult," IEEE transactions on neural networks, vol. 5, no. 2, pp. 157- 166, 1994.
[12] S. Selvin, R. Vinayakumar, E. A. Gopalkrishnan, V. K. Menon and K. P. Soman, "Stock price prediction using LSTM, RNN and CNN-sliding window model," in International Conference on Advances in Computing, Communications and Informatics, 2017.
[13] J. Schmidhuber, D. Wierstra, M. Gagliolo and F. Gomez, "Training recurrent networks by evolino," Neural computation, vol. 19, no. 3, pp. 757-779, 2007.
[14] L. Pasa and A. Sperduti, "Pretraining of recurrent neural networks via linear autoencoders," in Advances in Neural Information Processing Systems, NIPS, 2014, pp. 3572-3580.
[15] J.Chen and N. S. Chaudhari, "Segmented-memory recurrent neural networks," IEEE transactions on neural networks, vol. 20, no. 8, pp. 1267-1280, 2009.
[16] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural computation, vol. 9, no. 8, pp. 1735-1780, 1997.
[17] R. S. Sutton and A. G. Barto, Reinforcement learning: An introduction, MIT Press, 2018.
[18] F. A. Gers and J. Schmidhuber, "Recurrent nets that time and count," in Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000.
[19] https://colah.github.io/posts/2015-08-Understanding-LSTMs/