Check for updates

Stock Price Prediction Using Artificial Intelligence and Neural Networks

Kunal Chawla [1] , Dushant Singh [2] , Parth [3] , Aditya Raj [4]

1. Bachelor's in Engineering in Computer Science with Specialization in Artificial Intelligence & Machine Learning, Apex Institute of Technology, Chandigarh University, Gharuan, Punjab, India ,

2. Bachelor's in Engineering in Computer Science with Specialization in Artificial Intelligence & Machine Learning, Apex Institute of Technology, Chandigarh University, Gharuan, Punjab,, India ,

3. .Bachelor's in Engineering in Computer Science with Specialization in Artificial Intelligence & Machine Learning, Apex Institute of Technology, Chandigarh University, Gharuan, Punjab, India ,

4. Bachelor's in Engineering in Computer Science with Specialization in Artificial Intelligence & Machine Learning, Apex Institute of Technology, Chandigarh University, Gharuan, Punjab, India

Abstract: Predicting the stock prices is a very complex task, and to predict an almost accurate stock price,we need a robust and accurate algorithm which can analyze and compute the longer-term share prices.Several researcher's equally in the world and different industries have been very interested in the stockmarket. Stock processes are correlated within the nature of the market and that is why it is difficult topredict the share price. This project aims at processing and analyzing huge volumes of data (live data)and running comprehensive algorithms on the dataset. The purpose of the paper is to understand theshortcomings of the current prediction algorithms and to provide a method using neural networks andartificial intelligence through which we can predict the shared values with accuracy.By using the proposed method, anyone can monitor the preferred stock in real-time and can invest in thestock to make the most money by buying a large number of shares at the cheapest price and sellingthem at the highest price..

Keywords: stock price prediction, artificial intelligence, neural networks, algorithm, share prices, researcher, stock market, data analysis, shortcomings, prediction algorithms

- X -

INTRODUCTION

In present days, people are very interested in investing their money in the stock market to earn more in a short period. The cornerstone of any corporate portfolio is their stock, which can be bought privately. In order to stop any further illicit actions, any such transactions must be subject to established legal standards. "A stock is a type of security that specifies ownership in the issuing corporation." The stocks have survived the wrath of time and it can be brought to the public by using a number of ways.

Stock market trading is an extremely tough and complex system where people will either gain or lose fortune and also which require more experience and practice. In this work, an attempt is made to predict the stock prices by using Recurrent Neural Network and advanced machine learning concepts such as LSTM (Long Short Term Memory). This model considers the historical equity share price of a company and applies the above stated concepts in order to get the results. The different features of stock market shares are Opening price, day High, Close price, Date of trading, day Low, Total Trade Quantity and Total Trade Turnover. The proposed model fetches live and historical data from Yahoo Finance and after cleaning and transforming the data it is fetched into the keras model to give us the required result. The

keras model which is created contains five hidden layers of neural network to calculate the patterns in the stocks prices and it uses the 'relu' activation function to provide a threshold above which the values of the calculated weights is passed on to the next layer of neural network. We then create a summary table of the stocks based on the output and plot the values on the chart to showcase the predictions made by the model. Based on the calculations, we summarize that the model showcases the output in the graphical form with an accuracy of 80%-85% . This will enable the user to make better data driven decisions while investing into a stock.

LITERATURE SURVEY

The Dutch East India Co. was the first company to first ever issue stocks to the public. After this historical exchange the future of stocks changed forever. The stock market of a particular company is considered to be an accurate reflection of its economic effectiveness. The Stock Market is ever-changing. The prices of the stock keep on changing on a regular basis and they are prominently affected by the flow of finances throughout various economic domains and companies. The stock market is dependent on the demand and supply curve which is the backbone of economics. An increase in the demand of a particular stock; the price of the stock increases significantly and if the company is going through any economic crisis or recession the price of the stock goes down. Even with the fluctuations in the stock market the investments are made and then it becomes very necessary to predict the correct value of the stocks in the upcoming future in order to resist any further loss of fortune.

A number of machine learning algorithms are used in recent times to find out predictions for the share prices which are discussed in Table 1. The various established traditional approaches to stock price prediction are highlighted by this study. Furthermore, it addresses recent applications of machine learning methods as well as their advantages and disadvantages.

Figure 1: Predictive Analysis Process

Table 1: Comparative study of different approaches used for stock price prediction

S.NO.	TECHNIQUES USED	POSITIVES	LIMITATIONS
1.	Stock Market Prediction using Simple Linear Regression. [1]International Conference on Electronics, Communication and Aerospace Technology (ICECA 2017)	1. It forecasts the behavior of the TCS dataset using a linear regression algorithm, and the performance is then evaluated against the other methods. 2. The approach adopted contains methods for real-world applications, such as gathering and examining a vast dataset.	1. The model is only running on a single stock set which creates a degree of short-sightedness for the whole evaluation process. 2. Stock price prediction using linear regression is very less accurate.
2.	Stock Market Prediction using SVM and PCA. [2]2015 19th International Conference on System Theory, Control and Computing (ICSTCC), October 14-16, Cheile Gradistei, Romania	1. This project makes the use of SVM and Principal component analysis. 2. This model is quite accurate as it compares simple methods from SVM and evolves them to GASVM and PCASVM.	1. The complexity of the defined method is very high. 2. The paper does not provide any type of study to evaluate the developed model
3.	Stock market Prediction using Artificial Neural Networks. [3]2017 International Conference on Technical Advancements in Computers and Communications.	1. Uses sentiment analysis to produce more accurate results from the social media site 2. Model can be used in stock market, finance, auditing and business strategies	1. It does not provide the accuracy of the prediction. 2. There is not visual outcome to analyze the prediction.
4.	Stock Market prediction using Hidden Markov Model. [4]2013 IEEE International Conference on Granular Computing (GrC).	1. Decides which pattern matching sequences to use to produce a probabilistic approach 2. Avoids factor selection, combination and transformation. 3. Uses Hidden Markov Model to predict the stock prices.	1. The HMM model is only accurate for a very short period of time. 2. Module is not tested for a sample other than the training dataset.

5.	Stock Market Prediction using Sentiment Analysis: Hybrid Approach. [5]International Conference on Computing, Communication and Automation (ICCCA2016)	1. Uses both sentiment analysis and clustering to get accurate results 2. Hybrid model merges the positive factors of both the methods	1. No visual representation of prediction is depicted in the model 2. No regression techniques are applied 3. Does not give a model to compare fixed values.
6.	Multiple Linear Regression Model to predict Audit Opinions. [6]2009 ISECS International Colloquium on Computing, Communication, Control, and Management	1. Used those financial factors that have a vital impact on audit opinions. 2. Used Mean Squares and best fit line for prediction	1. This is based on only Shanghai and Shenzhen stock markets of 2007. 2. Does not clarify clean audit opinions
7.	Stock Market Prediction Model using ANN and Back Propagation. [9]ICCCNT 12 26th _28th July 2012, Coimbatore, India	1. Uses artificial neural network to give an overall prediction accuracy of 63% 2. Uses back propagation algorithm for training the model	1. This proposed model's real and empirical correctness varies greatly.

INFERENCE FROM RELATED WORK

As shown in Table 1, the present day studies are only focusing over the singularity of linear and multiple regression; which is not providing accurate results. The results are not accurate because Linear Regression is a Supervised learning algorithm in which we presume that the Independent Variable(Y) and Dependent Variable (X) are correlated and the best fit line among them is always linear. However, this is not the case while we are predicting the stock market as the nature and price of the stock changes every single minute. The present studies are only focusing over a single stock value and are training the model using a single stock value which affects the accuracy for other stocks. Therefore, it is pretty safe to conclude that there may be some other more accurate ways to predict the stock values.

PROPOSED WORK

In this paper, we are trying to predict the stock market trend by using algorithms of artificial intelligence and neural networks. As shown in Fig 2. The data is fetched from the online Yahoo finance and then pre-processing is done by removing any type of errors which are present in the dataset. The data frame then goes under the process of Feature Engineering where some of the features are selected over which the model is trained. Then the transformed data is sent into the artificial neural network with five hidden layers and then an output is generated which is then mapped over the graphs using the matplotlib. The trained model is then saved and is applied over the front-end to display the results over a web app.
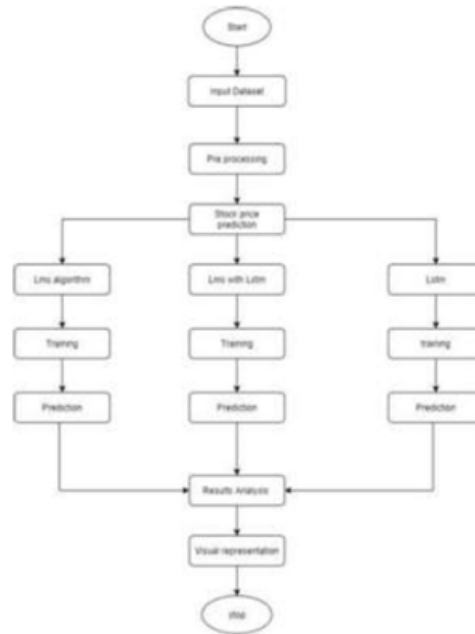
Figure 2: Data Flow Diagram

- **Technologies Used**

To provide accurate results we have used the most prominent yet user friendly language Python. The model is built using Python and the front end is made by using the python library streamlit. For web hosting we have used the streamlit web hosting services.

- **Libraries Used**

1. **Numpy:** This Python library offers multidimensional array objects and quick array operations, including mathematical, logical, shape manipulation, sorting,selecting, I/O, basic statistical operations, random simulation and much more.

2. **Pandas_Data-reader:** It is a python package which allows the user to create a pandas dataframe object by using various online data sources such as Yahoo Finance, Google Finance, Morningstar, FRED, World Bank.

3. **Keras:** This is the high level API of Tensorflow which has a highly productive interface for solving complex machine learning problems , with a major focus on modern deep learning. It also provides essential abstractions and building blocks for developing and shipping machine learning solutions.

4. **Tensorflow:** This is an end-to-end, open source machine learning platform. It was originally developed by Google and it comes with great support for machine learning and deep learning and the flexible numerical computation core is used for many scientific domains.

5. **Matplotlib:** It is also a python library which is majorly used to visualize our outcomes over a graph for better understanding of the data and the output

6. **Streamlit:** It is a very promising open-source library to make beautiful machine learning and deep learning web applications. It is a Python based library. It is the easiest way especially for people with not front-end development knowledge.

- **Algorithm/ Concepts Used**

1. LSTM (Long Short Term Memory): LSTM is the advanced version of RNN (Recurrent Neural Network) where the information belonging to the previous state is stored and present. These are very different from the regular RNN because they involve long dependencies but RNN works on finding the relationship between current information. The main advantage of using this algorithm is that it is dependent over large amounts of data As stock market analysis requires a lot of historical data the LSTM model also regulates the error by giving aids to the Recurrent Neural Network by retaining the information for the older stages which is making the prediction more accurate. The gradients with respect to the weight matrix may become very small and can degrade the learning rate of the model. The stock market has a vast amount of historical data and the LSTM model predicts the prices of the stock more accurately within a short-time period with greater accuracy as compared to any other model.
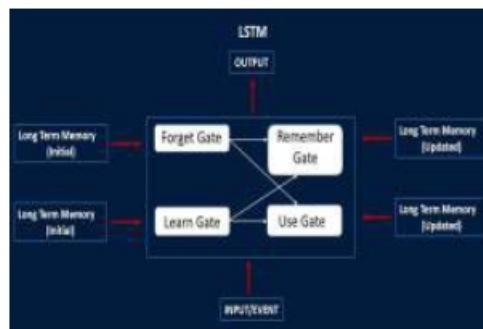
The above data flow diagram correctly depicts the total amount of LSTM layers utilized by the model for getting the prediction between 80%-85%. Dropout is a regularization method where the input and recurrent connections to LSTM are taken as probabilistic and is excluded from activation and weight updates while training a network. In the model we have utilized different dropout values at each LSTM layer for getting accurate predictions.
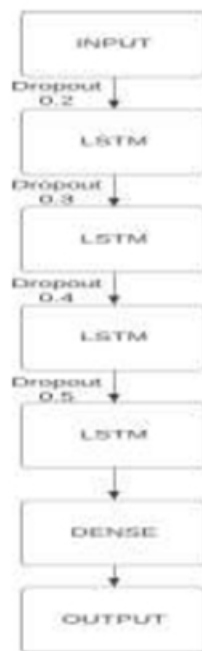
Fig 3: Data flow diagram of different layers and dropout values used in the model

2. Neural Network: Now the training data is created, now we need to make a model for the time-series prediction for this we require the Tensorflow framework of Python. A sequential type of model is created where each neural network layer is connected with each other and the training data is passed after which an output is generated and then it is passed on to the next layer. Finance is very non-linear and sometimes the price of the stock may be very random. Traditional time series models such as ARIMA and GARCH are only effective over static data. The main issue in using these models is their complexity of implementation in a live trading system, as there is no guarantee of stationary data.

- **Activation Function**

The majority of activation functions implemented in artificial neural networks (ANN) output gives modest values for small inputs and higher values for inputs that are greater than the threshold value. The activation function transmits the data from one neuron to another if the value is high enough. These are helpful because they give the neural network non-linearities, which enable the neural network to learn powerful operations quickly.

There are different types of activation functions such as sigmoid, relu, leaky relu and tanh. After comparing and implementing all the results of the above stated activation function we found out that for the proposed model; relu is providing the highest and efficient accuracy for the prediction of the results. Relu stands for rectified linear activation function or we can say a nonlinear function that will output the input directly if it is positive, or else the output will be zero. The mathematical expression of relu is $f(x) = max(0,x)$. For positive values, Relu behaves as a linear function; for negative values, it behaves as a non-linear activation function. Relu is basically simpler in computation as its derivative remains constant I.e.1 for all the positive inputs and this reduces the time taken by the model to calculate the derivative for each and every positive value. It is also capable of outputting a true zero value and the linear activation function is easier to

optimize and allow a very smooth flow in the model.

- **Analysis of Data**

Analysis of the given data is done using TABLEAU. Tableau is a tool that can be used to analyse data and create visualizations. Tableau allows its users to create interactive dashboards with rich visuals and explore data in the context of their business. Tableau comes with different types of visualization such as histograms, scatter plots, pie charts bubble charts etc. It also has advance calculation fields, data filtering options, we can create cross table. It is great tool for data analysis and data visualization.

Figure (4)

Bubble chart also known as text chart is used to show the volume of the item sets. It is an extension of scatter plot. The above figure (4) shows the volume of each stock. We are filtering based on certain criteria as: "volume of the stock" given minimum value of volume i.e., 180000000. Here we have analysed the Stocks of four big companies:

 (A) Amazon

In the above figure(5) we are analyzing the Amazon stock by doing the comparison between volume and price over the years, using a 'dual axis chart'. The cross table chart is used to display the open price, close price, low, high price, and volume of the stock over the years. In the dual line chart, the 'blue' coloured line is showing the trend of the price of stock whereas 'orange' coloured line is showing the trend of volume of stock. In 2010, the initial volume of the stock was 1852790297.254 and the opening high price was 35,509. As the trends displayed in the graph with year, as the volume decreases till year 2013, there is a hike in the price. In 2014, the volume and price has proportionally increased. And last in 2016, the volume has dropped from 1852790297 to 1037805 and price has increased from 35,509 to 177,863. Thus we can say that with time as volume decreases, the high price of the stock increases.
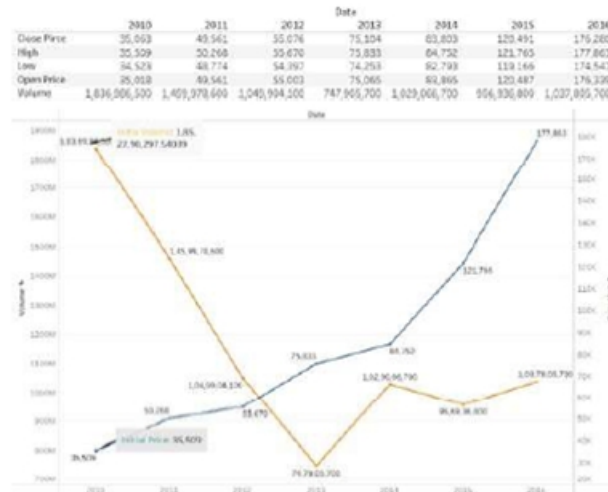
Figure (5) Amazon

(B) Apple

In the above figure (6) we are analyzing 'APPLE' stock by doing the comparison between volume and price over the years, using 'dual axis. The cross table chart is used to display the open price, close price, low, high price, and volume of the stock over the years. In the dual line chart, in year 2010, the initial volume of the stock was 39756474201.74 and the opening high price was 66,117. As the trends displayed in the graph with year, as the volume decreases till year 2011 and increases for 1 year and after 2013, huge drop in volume can be observed till year 2016 i.e. from 39756474201 to 9680671300. Similarly, the price of the stock increases till year 2012 i.e. from 66,117 to 145,456. And after 2012, there is huge drop in price from 145,456 to 26,568. Here the both price and volume have decreased

Figure (6) Apple

PROJECTILE ANALYSIS

This work is an attempt to showcase the abilities of artificial neural networks and artificial intelligence in

the field of stock market price prediction. The model which is developed has certain steps in order to provide us with accurate results. Here the input is the historical data of the stocks and the expected output is the prediction of stock price variation.

Steps in stock price prediction:

1. **Data collection:** Data collection is the process of collecting all the necessary data on which the model will be created, tested and trained. Collecting data which is unbiased is very important for accurate predictions. In this case all the data was collected from the Yahoo Finance website and the model was trained and tested over it.

2. **Data pre-processing and cleaning:** Data preprocessing is the process of transforming the data into required information. The data which is collected may have certain flaws or missing values in it, so we need to remove those discrepancies from the data for a smooth process. If the data contains missing values then it can be recovered by using imputation techniques .

3. **Data Normalization:** Data is then needed to be standardized in order to get better accuracy by ensuring that all the features are in similar range.

4. **Feature Extraction:** This step involves searching in the space of possible features. We then pick the subset that is optimal or near-optimal with respect to some objective function. This is majorly done to prevent any kind of overfitting or under fitting in the data set.

5. **Splitting Data:** Splitting of data is done in order to rain the model on a particular set of data and test it over the other set of data in order to prevent any kind of discrepancies. So the model is generally splitted into 70% of training data and 30% of training data.

6. **Building RNN model:** This is a very crucial step as the major prediction will be done by the recurrent neural network model. In this model we initialize RNN by using the sequential repressor and then we have adding the LSTM layer and dropout regularization is used for the Removal of any undesirable values will improve model effectiveness. Five LSTM layers with different dropout values are fed to the model in order to predict the accurate price of the stock.

7. **Saving the Model:** The keras model is then saved so that it can be applied over the live data to provide live results to the user.

8. **Testing the model:** The testing part is where we utilize the other 30% part of the data to test if our model is predicting accurate results and what is the accuracy we can expect from our model.

9. **Create Web Application:** Now after creating, training and testing the data, comes the part of representing the data. For this we use the python based library streamlit to create a web application in which there is a input section for entering the stock ticker and data is represented from 2010-2022 and afterwards the data is fed into the model in order to bring out the result of the predicted value which is then compared with the original value with the predicted value over a graph.
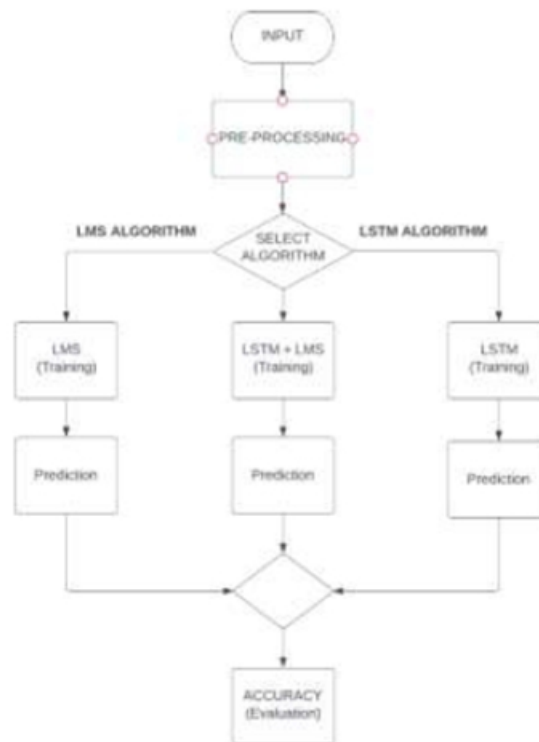
Figure 7: Steps of Implementing Stock Price Prediction

RESULTS & DISCUSSION

The proposed solution's implementation i.e. RNN model using python3 which predicts the future trends of the Apple Inc., stock is mentioned below in Fig 5. The below visualization shows the actual vs the predicted price of the Apple stocks. After reviewing the accuracy of the model for the Apple stock, we get that the model is having and accuracy of 87.5%, which is far better than any other stock price prediction algorithm until date.
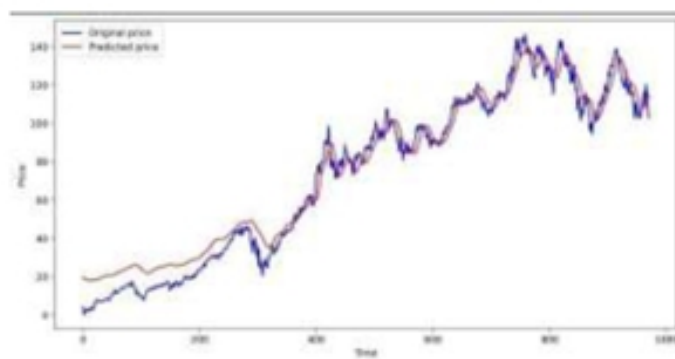

Figure 8: Result of APPLE stock Price Original vs Predicted

Other stock prices predictions can also be generated by using the same model as it is fetching the data of the specified Stock Ticker in real time such as in Fig 6. The stock price prediction of Tesla is showcased in which the model has plotted the actual vs predicted price of the stock by using the same RNN model. Here

the data is showcased from Year 2010 to 2022 and then the Closing vs Time chart is plotted and afterwards the RNN model predicts the price of the stock with 82.2% accuracy.

	High	Low	Open	Close	Volume	Adj Close
count	3,112.0000	3,112.0000	3,112.0000	3,112.0000	3,112.0000	3,112.0000
mean	58.7999	56.1041	57.5090	57.4809	93,277,742.2237	57.4809
std	97.6082	92.9522	95.4097	95.3028	81,999,591.0860	95.3028
min	1.1087	0.9987	1.0760	1.0533	1,777,500.0000	1.0533
25%	8.5517	8.1258	8.3957	8.3500	41,863,500.0000	8.3500
50%	16.3357	15.7913	16.1000	16.1167	75,410,250.0000	16.1167
75%	24.3520	23.4953	23.9225	23.8597	117,083,775.0000	23.8597
max	414.4967	405.6667	411.4700	409.9700	914,082,000.0000	409.9700

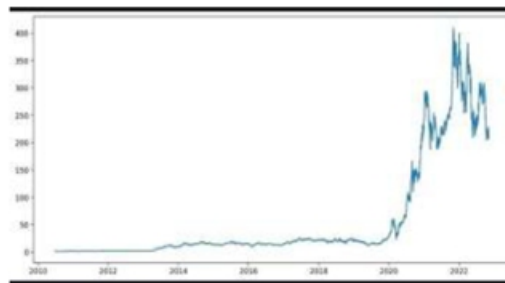**Figure 9: Actual Data from 2010 to 2022**

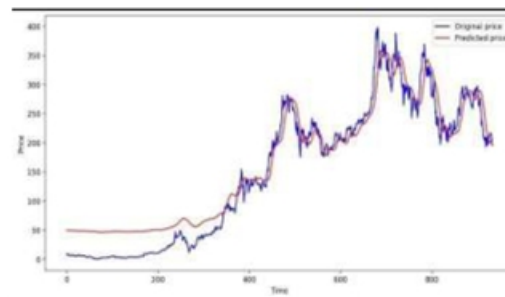
Figure 10: Closing v/s Time Chart

Figure 11: Actual vs Predicted Chart of TSLA stock

FUTURE WORK

Stock prices are very complex and they rely on various factors and the Intangible factors like human sentiments, social media influence, and brand reputation are not being used in the model but they may have an impact on the market. There may be accurate results. Another feature of live chart patterns is missing from the model , which can be made by applying an algorithm to predict the prices on every minute and which can be showcased over the graph. The ability to push alert notifications can be added in future to the site in order to tell the user when the price of a particular stock will get up so that they can invest on that particular stock at the most optimum time. New embedded learning can be applied and more LSTM layers can be added to provide better accuracy of the model by using other models along with LSTM.

CONCLUSION

This study aims to assist stock market brokers and investors in making informed and prudent financial decisions. Prediction has a huge range of applications in the stock market, but because of the always shifting stock market facade, this is a highly difficult task. This paper presented the positives and shortcomings of all the other methods, which are used for the stock price prediction and also presented a more accurate method to find out the future price of the stock.

We utilized the complex concepts of artificial neural networks and artificial intelligence to find out a more accurate solution to the problem. The proposed model is analyzed to provide an accuracy for every stock between 80%-85%. The predictions made by the LSTM model using the dropout results in increased accuracy and accurate predictions in less amount of time. This model can be utilized to predict the stock price of any stock, which is listed over the Yahoo Finance website.

References

1. Hiba Sadia, Aditya Sharma, Adarrsh Paul, SarmisthaPadhi, Saurav Sanyal- "Stock Market Prediction Using Machine Learning Algorithms", IJEAT, 2019.

2. Raut Sushrut Deepak, Shinde Isha Uday, Dr. D. Malathi, "Machine Learning Approach In Stock Market Prediction", IJPAM 2017

3. M. Roondiwala, H. Patel and S. Varma, "Predicting stock prices using LSTM," International Journal of Science and Research (IJSR), vol. 6, no. 4, pp. 1754- 1756, 2017.

4. M. S. Hegde, G. Krishna and R. Srinath, "An Ensemble Stock Predictor and Recommender System," 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Bangalore, 2018, pp. 1981-1985.

5. T. Kim and H. Y. Kim, "Forecasting stock prices with a feature fusion LSTM-CNN model using different representations of the same data," PloS one, vol. 14, no. 2, p. e0212320, April 2019.

6. 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.

7. Loke.K.S. "Impact Of Financial Ratios And Technical Analysis On Stock Price Prediction Using Random Forests", IEEE, 2017

8. Xi Zhang1, Siyu Qu1, Jieyun Huang1, Binxing Fang1, Philip Yu2, "Stock Market Prediction via Multi-Source Multiple Instance Learning." IEEE 2018.

9. Tao Xing, Yuan Sun, Qian Wang, Guo Yu. "The Analysis and Prediction of Stock Price", 2013 IEEE International Conference on Granular Computing (GrC), ISBN: 978-1-4799-1282-7, DOI:10.1109/GrC.2013.6740438.

10. Jordan Prosky, Andrew Tan, Xingyou Song, Micael Zhao. "Sentiment Predictability for Stocks" - arXiv:1712.05785v2 [cs.CL], 18 Jan 2018.

11. X. Shao, D. Ma, Y. Liu, Q. Yin. "Short-term forecast of stock price of multi-branch LSTM based on K-means", 2017 4th International Conference on Systems and Informatics (ICSAI).

12. Stock Market Prediction using Simple Linear Regression. [1]International Conference on Electronics, Communication and Aerospace Technology (ICECA 2017)

13. Stock Market Prediction using SVM and PCA. [2]2015 19th International Conference on System Theory, Control and Computing (ICSTCC), October 14-16, Cheile Gradistei, Romania

14. Stock market Prediction using Artificial Neural Networks. [3]2017 International Conference on Technical Advancements in Computers and Communications

15. Stock Market prediction using Hidden Markov Model. [4]2013 IEEE International Conference on Granular Computing (GrC).

16. Stock Market Prediction using Sentiment Analysis: Hybrid Approach. [5]International Conference on Computing, Communication and Automation (ICCCA2016)

17. Stock Market Prediction Model using ANN and Back Propagation. [9]ICCCNT'12 26th _28th July 2012, Coimbatore, India

18. Multiple Linear Regression Model to predict Audit Opinions. [10]2009 ISECS International Colloquium on Computing, Communication, Control, and Management

19. Real-Time Stock Prediction using Neural Network. [8]2018 8th International Conference on Cloud Computing, Data Science & Engineering (Confluence)