Review on Stock Market Prediction

[1]**Dr. Archana Dehankar,** [2]**Ayush Itankar,** [3]**Abir Meshram,** [4]**Nayan Bhople,** [5]**Bhushan Patle**

[1]Assistant professor, Department of Computer Technology, Priyadarshini College of Engineering, Nagpur, India
[2,3,4,5]Student, Department of Computer Technology, Priyadarshini College of Engineering, Nagpur, India

Abstract - **Stock trading is one of the most important practices in the financial world. The act of attempting to forecast the future value of a stock or other financial instrument traded on a financial exchange is known as stock market prediction. The majority of stockbrokers use technical and fundamental analysis, as well as time series analysis, when making stock predictions. Python is the programming language used to use machine learning to forecast the stock market. In this paper, we propose a Machine Learning (ML) method that will be trained using publicly accessible stock data to obtain intelligence, and then use that intelligence to make an accurate prediction. The paper focuses on the use of Regression and LSTM based Machine learning to predict stock values. Factors considered are open; close, low, high and volume. The dataset should be as precise as possible, because even minor changes in the data might result in large changes in the results. In this paper, supervised machine learning is employed on a data set obtained from Alpha vantage api.**

Keywords: RNN (recurrent neural network), machine learning, financial times series forecasting, stock market prediction, linear regression.

I. INTRODUCTION

A good stock forecast will result in huge profits for both the seller and the broker. It is often mentioned that prediction is chaotic, meaning that it can be predicted by carefully studying the background of the related stock market. Machine learning is a good way to describe these types of processes. It forecasts a market value that is similar to the tangible value, which increases accuracy. Because of its effective and precise measurements, the application of machine learning to the field of stock prediction has piqued the interest of many researchers.

The dataset used in machine learning is critical. Since even small changes in the data can result in massive changes in the outcome, the dataset should be as accurate as possible. On a dataset obtained from Alpha vantage, supervised machine learning is used in this paper. The following five variables make up this dataset: open, low, high, and volume. Different bid prices for the stock at different times with nearly direct names are known as open, high, low, and close. The number of shares that passed from one owner to another

during the time span is referred to as the volume. After that, the model is put to the test with the test results.

Regression and LSTM models are engaged for this conjecture separately. Regression involves minimizing error and LSTM contributes to remembering the data and results for the long run. Finally, the graphs for the fluctuation of prices with the dates (in case of Regression based model) and between actual and predicted price (for the LSTM based model) are plotted.

II. LITERATURE SURVEY

[1] Survey on stock market prediction using machine learning techniques

The objective is to predict the market performance with the help of an artificial neural network. The techniques of artificial neural networks classify the stock in mainly three categories that is buy, hold and sell, supported previous data. It's observed that the logistic regression model is employed by every individual to predict a stock in an exceedingly better way.

[2] Survey on stock market prediction using machine learning algorithms

The objective is to predict the stock movement with good accuracy. The realm of interest is that the dataset of the stock prices from past years. The raw dataset must be pre-processed for data analysis. After pre-processing the information, we are going to use machine learning techniques like random forest and support vector machines on the dataset to come up with the outcomes.

[3] Survey on predicting stock prices using LSTM forecasting the stock prices has always been a difficult task for the analysts II

The investors are highly curious about the stock prediction. For a successful investment, many investors have an interest in knowing the long-run condition of the stock market. The prediction models for the stock market helps the investors and also the analysts by providing the long run information of the stock market. Recurrent neural networks (RNN) and Long Short-Term Memory (LSTM) is the machine learning approaches used for stock price prediction.

[4] Survey on stock trend prediction using regression Analysis

A data mining approach various organizations are collecting data, building large data warehouses to store the collected data. Discovering the knowledge out of the collected data is done by employing a technique called data mining. Data mining software tool is employed to extract values of the variables from the dataset to predict the long-run values of other variables with the utilization of time series data.

[5] Survey on stock market prediction using machine learning

Developing an application for analyzing and predicting stock market prices increases the investor's interest in stock markets. First, we have to analyze the present and emerging methods of stock price prediction. The different approaches are fundamental analysis, technical analysis, and also the application of machine learning. Fundamental analysis and machine learning are accustomed to guide an investor's decisions. Whereas the technical analysis methodology provides limited useful information.

[6] Survey on stock market prediction: using historical data analysis

The stock market plays a vital role in business and finance. Fundamental analysis is depleted using social media data with the help of Sentimental Analysis. Social media data includes a high impact because of its increased usage, and it helps predict the trend of the stock market. Technical analysis is depleted by applying machine learning algorithms on historical data of stock prices. The tactic usually involves gathering various social media data, news to extract sentiments expressed by individuals.

III. METHODS FOR STOCK MOVEMENT PREDICTION

A. Methods for Stock Movement Prediction the Support Vector Machine (SVM) has long been recognized as being able to efficiently handle high dimensional data and has been shown to perform well on. Therefore, we chose the SVM with the linear kernel as the prediction model. To assess the effectiveness of sentiment analysis on the message boards, six sets of features are designed. The first one used only the historical prices. The other methods incorporated the mood information into the prediction model. All the feature values were scaled into $[-1, 1]$.

B. Price only in this method, only historical prices are used to predict the stock movement. The purpose of this method is to investigate whether there are patterns in the history of the

stock or not. In addition, this model was used as a baseline to evaluate whether integration of the sentiments is effective by comparing with other sentiment models. Features used for the training of SVM are $pricet-1$ and $pricet-2$ which are the price movements (up, down) at the transaction dates $t-1$ and $t-2$, respectively.

C. Human sentiment in addition to historical prices, this model integrated the sentiments annotated by humans into the prediction model. The Message Board dataset, the users explicitly select a sentiment label with their posts. These sentiment labels are "strong buy", "buy", "hold", "sell" and "strong sell". Instead of using all the messages, we tried to use only the messages with annotated sentiments by the users, and discard the other messages. From these messages, we used only the explicit sentiment and removed other in-formation such as message content. The purpose of this method is how mood annotated by humans can be used to predict the stock. Because the sentiments are annotated by humans, this feature is one of the strongest features for stock prediction.

D. Sentiment classification to utilize the remaining 84.4% of the messages without the explicit sentiments, we tried to build a model to extract the sentiments for those messages. A classification model was trained from the messages with annotated sentiments on the training dataset. Then it was used to classify the remaining messages into five classes (Strong Buy, Buy, Hold, Sell, and Strong Sell).

E. LDA-based method in this model, we consider each message as a mixture of hidden topics. Latent Dirichlet Allocation (LDA) (Blei, Ng, & Jordan, 2003) is a generative probabilistic model of a corpus. The basic idea is that documents are represented as random mixtures over latent topics, where each topic is characterized by a distribution over words. Therefore, we choose the LDA as a simple topic model to discover these hidden topics.

F. JST-based method the opinion is often expressed on a topic or aspect. When people post a message on social media to express their opinion for a given stock, they tend to talk about their opinions for a certain topic such as profit and dividend. Based on pairs of topic-sentiments, they would think that the future price of that stock goes up or down. From that intuition, we propose a new feature topic-sentiment for the stock prediction model. To extract pairs of topic-sentiment, we tried to use two kinds of models.

G. Aspect-based sentiment Instead of considering the mixtures of hidden topics and sentiments as in the previous model, in this model the mixtures are not hidden. Each message is represented as a list of topics and their corresponding sentiment values. In our proposed method, the topic is the consecutive nouns in the sentence.

IV. CONCLUSION

Stock investing has attracted the interest of many investors around the world. However, making a decision is a difficult task as many things are involved. By investing successfully, investors are eager to predict the future of the stock market. Even the slightest improvement in performance can be enormous. A good forecasting system will help investors make investments more accurate and more profitable by providing supporting information such as future stock price guidance. In addition to historical prices, other related information could affect prices such as politics, economic growth, financial matters and the atmosphere on social media. Numerous studies have proven that emotional analysis has a significant impact on future prices. Therefore, the combination of technical and basic analysis can produce very good predictions.

This paper proposes RNN based on LSTM built to forecast future values. The testing result will confirm our model is capable of tracing the evolution of opening prices. For our future work we will try to find the best sets for bout data length and number of training epochs that beater suit our assets and maximize our predictions accuracy.

REFERENCES

[1] Kraus and Feuerriegel [51],German ad hoc announcements, Transfer learning with RNN and LSTM, Word embedding (em),Polarity score, Direction accuracy of nominal return, Result- RNN 0.552LSTM 0.576 LSTM-em0.578.

[2] K. Raza, "Prediction of Stock Market performance by using machine learning techniques," 2017 International Conference on Innovations in Electrical Engineering and Computational Technologies (ICIEECT), Karachi, 2017, pp. 1-1.

[3] Huynh et al. [47], S&P 500 index, financial news, method(BGRU),Word embedding, real value vector, measure used(Prediction accuracy),result_59.98%.

[4] M. Billah, S. Waheed and A. Hanifa, "Stock market prediction using an improved training algorithm of the neural network," 2016 2nd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE), Rajshahi, 2016, pp. 1-4.

[5] H. L. Siew and M. J. Nordin, "Regression techniques for the prediction of stock price trend," 2012 International Conference on Statistics in Science, Business and Engineering (ICSBE), Langkawi, 2012, pp. 1-5.

[6] K. V. Sujatha and S. M. Sundaram, "Stock index prediction using regression and neural network models under non-normal conditions," INTERACT-2010, Chennai, 2010, pp. 59- 63.

[7] S. Liu, G. Liao, and Y. Ding, "Stock transaction prediction modeling and analysis based on LSTM," 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA), Wuhan, 2018, pp. 2787-2790.

[8] T. Gao, Y. Chai, and Y. Liu, "Applying long short term memory neural networks for predicting stock closing price," 2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS), Beijing, 2017, pp. 575-578.

[9] K. A. Althelaya, E. M. El-Alfy, and S. Mohammed, "Evaluation of bidirectional LSTM for short-and long term stock market prediction," 2018 9th International Conference on Information and Communication Systems (ICICS), Irbid, 2018, pp. 151-156.

[10] Li et al. [52], CSI300 index data, investors' forum posts, methods used-LSTM, NB, selection of textual data manually labelled sentiment by experts, measure used -Direction accuracy, result - 87.86%.

****