Stock Price Prediction

Amarpreet Singh
AIT-CSE
Chandigarh University
India
amar07sanu@gmail.com

Mohit Kumar Daga
AIT-CSE
Chandigarh University
India
mrmohitdaga@gmail.com

Ms.Bhavna Nayyer
AIT-CSE
Chandigarh University
India
E15505@cuchd.in

***Abstract*—In this work, we forecast stock price using machine learning. Stock price movements can be predicted using machine learning. Stock price forecasting aims to assist investors in making more precise and knowledgeable financial decisions.To improve stock forecast accuracy and generate lucrative trades, we advise using a stock price prediction method that incorporates mathematical techniques, machine learning, and other outside factors.There are two distinct stock types. You may be familiar with the phrase "day trading," which refers to intraday trading. Interday traders typically hold equities for a few days, a few weeks, or even a few months.LSTMs are effective at addressing problems involving sequence prediction because they can remember past data. This is relevant to our situation since it might be challenging to predict a stock's future price without first knowing its past price. Even if it can be difficult to predict a stock's price exactly, we can create a model that will show whether it will rise or fall.The goal of stock market prediction is to forecast the value of a company's financial stocks in the future. The use of machine learning, which produces forecasts based on the values of current stock market indices by training on their prior values, is a recent trend in stock market prediction technologies. Multiple models are used by machine learning itself to facilitate and authenticate prediction. The research focuses on the use of machine learning techniques based on LSTM to forecast stock values.**

Close, high, low, open, regression, and volume are some of the keywords.

I. Introduction

1.1 Problem Definition

The study of the relationship between a stock's price and its fundamental value is known as stock price prediction. The technique of forecasting a stock's or index's future price is known as stock price prediction. It is a challenging subject to tackle, and many variables, such as macroeconomic and industry trends, historical performance, technical analysis, fundamental analysis, and more, go into creating such predictions. Stock price prediction is the practise of estimating a stock's future worth based on past price movement patterns. From the financial markets to academic research, stock price forecasting can be applied in a variety of scenarios.

1.2 Problem Overview

A statistical method known as a stock price prediction model analyses previous data on a company's stock price to forecast what the current stock price will be. Using historical correlations between observed and unseen variables as well as predictive models for these variables is the most popular method for creating stock price prediction models.The effectiveness of this strategy rests on two aspects: how effectively the prediction models work when used with fresh data, and how clearly the past associations can be connected to the future ones. As part of our study, data from the stock market is gathered and used to forecast stock prices. To produce forecasts, we will use computational techniques, statistical techniques, and market information.

1.3 Hardware Requirements

1. RAM: 4 GB 2. Capacity: 500 GB 3. A 2 GHz or higher CPU 4. 32-bit or 64-bit architecture

1.4 Software Specification

1. Data pre-processing, model training, and prediction are performed in Google Colab using Python 3.5. 2. Operating system: MAC OS, Linux-based OS, Windows 7 and higher

II. Related Work

It was noted from the literature review that machine learning techniques are being applied to stock market forecasting extensively around the world. Comparing machine learning approaches to modern prediction methodologies, machine learning techniques are proving to be significantly faster and more accurate. In this field, significant work has been done all around the world. Which the work of M. Usmani, S. H. Adil, K. Raza, and S. S. A. Ali [1] and that of K. Raza is a witness to [2] who have examined how machine learning techniques are used and discussed recent developments in this area. Using deep neural network approaches, H. Gunduz, Z. Cataltepe, and Y. Yaslan [3] forecasted stock prices. Similar to this, M. Billah, S. Waheed, and A. Hanifa [4] proposed more enhancements to stock prediction using neural networks by employing a training technique that they independently created. K. V. Sujatha and S. M. Sundaram [6] developed enlightening methods for dealing with abnormal circumstances that frequently occur during system operation and might disrupt operations or result in incorrect forecasts. Similar research was done by Liu, G. Liao, and Y. Ding [7] who also created a model for applying Stock prediction using LSTM has a lot of room for accuracy improvements. K. A. Althelaya, E. M. El-Alfy, and S. Mohammed [9] made additional contributions to the field by setting up simulations and tests to evaluate if deep learning techniques may be used to predict stock values.

Identify applicable funding agency here. If none, delete this.

III. LITERATURE SURVEY

The stock price prediction project's literature review is an analysis of the previously conducted research on the subject by various academics. The goal of this study is to develop a technique for forecasting stock prices so that investors can choose their investments with more knowledge. The study will also incorporate findings from scholarly investigations on related topics that were conducted in the past.

2.1 Existing System Project on Stock Prediction: This project is an overview of the research on techniques for predicting stock prices. The objective of this project is to compile an assessment of the present body of knowledge about stock price prediction techniques and to identify potential topics for further study. We will review current approaches, assess their efficacy, and determine whether they are worthwhile for use in actual practise. We will also offer suggestions for further study. The stock price prediction project is an investigation into the effects of alleged "black swans" on the dynamics of stock prices. Black swans are uncommon occurrences that cannot be predicted by current models, but when they do, they can have significant repercussions. In this study, we analyse historical data to investigate how black swans effect stock prices.

Project: Current System of Stock Price Prediction Our group is presently engaged in a fresh endeavour: forecasting the stock price of a business. We have already created an algorithm that, based on a few variables, can forecast the price in the future. Our objective is to improve the usability and accuracy of this project. The model is constructed using genetic algorithms and neural networks, with several layers of genetic algorithms and multiple layers of neural networks. The model was evaluated for three days on fresh data after being tested for ten days on historical data. It was 85

2.2 Proposed System Our methodology is a fresh method of forecasting stock prices. We're attempting to forecast the stock's future price. We want to be able to inform you of market developments now and how they will effect your investment 30 days from now. Our algorithm operates by examining historical information regarding a company's stock price and producing forecasts based on that information. Our projections will be more accurate the more information we know.

To determine what influences a company's stock price, our system will employ machine learning approaches (differential equations and convex optimisation). This enables us to forecast future outcomes based on current events, which aids in the development of a precise model of how investors act when purchasing a company's shares. The goal of the proposed stock price prediction project is to develop a system that can forecast a stock's price before it is made available to the general public. A model known as LSTM will be used to create the system.

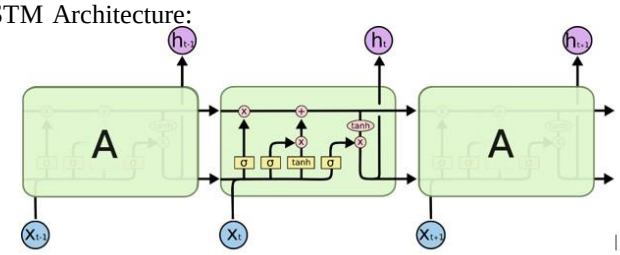
Network of long-term memory:

A specific type of recurrent neural network (RNN) is the long short-term memory network (LSTM).

LSTM operation:

The special network structure known as LSTM is composed of three "gate" components. An LSTM unit has three gates: an input gate, a forgetting gate, and an output gate. Information can be chosen by rules when it enters the LSTM network.The forgetting gate will remove any data that doesn't follow the algorithm, leaving only the information that does. The historical data used in this study's experiments were gathered from the Internet. The investigations made use of three data sets. An optimisation algorithm that requires less resources and has a faster rate of convergence must be found. used the embedded layer of the LSTM and the LSTM neural network with automated encoder. Gradients are kept from inflating and disappearing by using LSTM instead of RNN. In this project, MATLAB is used to minimise the input dimensions while Python is used to train the model. MySQL is used as a dataset for data storage and retrieval. The table of historical stock data includes data on the opening price, maximum price, lowest price, closing price, transaction date, volume, and other metrics.

LSTM Architecture:



IV. PROPOSED METHODOLOGY

Gather and Prepare Data: The first step is to gather a sizable data collection of the stocks for future predictions. You must get the dataset ready so that your machine learning algorithm may use it. This entails preparing the dataset beforehand, cleaning and removing uncertainties, and normalizing it so that it becomes handy for further operations.

The next phase is feature extraction, which comes after data preparation. This entails removing significant features from the datasets that can be efficiently used for predictions. Training: Using the labeled dataset, a machine learning model must be trained after the features have been extracted. The algorithms Decision Trees, Support Vector Machines, Naive Bayes, and Neural Networks are frequently used for stock price predictions. Here we have used Deep Learning model called as LSTM for our stock price predictions.

Validation: Using a validation set, you must verify the model's effectiveness after training. This ensures that the model can generalise to new data and does not overfit the training set. Testing is the last stage after validating the model's functionality. In order to assess the model's performance and

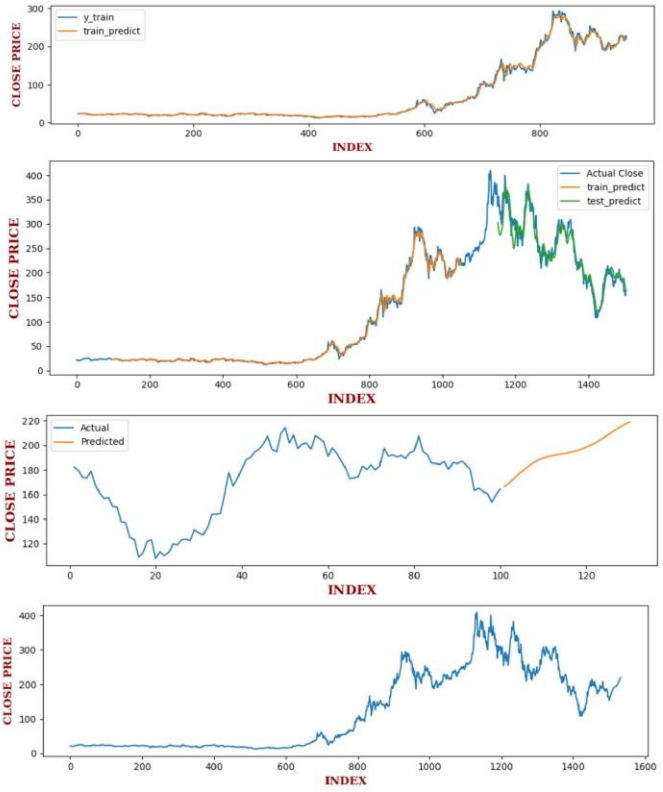
accuracy. You test it on a different test set in this phase. This process helps assess how well the model can predict on the new set of data.

Fine-tuning: You can adjust the model's parameters to increase accuracy if its performance isn't up to par. To enhance the performance of the model, you can experiment with various learning rates, number of epochs and optimisation strategies. Deployment: After adjusting, you may use the model to predict the future price of stock. Depending on the use case, this can be accomplished by connecting the model with a web application or a mobile app.

V. Experimental Result

The dataset used for training and testing the proposed system comes from Yahoo Finance. It is divided into training and testing sets, and after running through various models, produces the results shown below:

A. LSTM Based Model



The orange line represents the prediction, and the blue line represents the actual trend. These two lines' proximity reveals,

how effective the model based on LSTM is. When a significant amount of time has passed, the prediction comes close to the actual trend. A Train Score of 5.320054469 MSE (RMSE:

2.30651) and a Test Score of 18.91342393 MSE (RMSE: 4.348956) were obtained using the model. The accuracy that is reached will increase when the system is taught more and a larger dataset is used. More accurately than any other machine learning model, the LSTM Model was presented.

Fig.4 Plot comparing LSTM's Actual and Predicted Trend The orange line represents the prediction, and the blue line represents the actual trend. These two lines' proximity reveals,

VI. Dataset Description

The dataset that will be analysed was obtained from Yahoo Finance. About 1500 records of the necessary stock prices and other pertinent values were included in the dataset. The information showed the stock prices for each day of the year at specific intervals. It was divided into several parts, including date, symbol, open, close, low, high, and volume. The data for just one company was taken into account for simulation and analysis. The data was all contained in a csv file, which was read and then converted into a data frame using Python's Pandas package. By separating the data based on the symbol field, the data for a specific company was then taken from this.

Following this, the data was normalised using the Python sklearn module, and it was split into training and testing sets. 30

VII. Evaluation metrics used

Error Root Mean Square An addition to the mean squared error is the Root Mean Squared Error, or RMSE.The square root of the error is determined, which is significant since it means that the units of the RMSE and the goal value that is being forecasted are the same.For instance, if your target variable has the unit "dollars," the RMSE error score will also contain "dollars" and not "squared dollars" like the MSE.As a result, it would be typical to train a regression predictive model using MSE loss and then evaluate and report its performance using RMSE.Following are the steps to calculate the RMSE:RMSE is equal to sqrt(1/N*sum for i to N (yi - yhati)2) where yhat i is the i'th projected value, sqrt() is the square root function, and y1i is the i'th expected value in the dataset.The RMSE can be restated as RMSE = sqrt(MSE) in terms of the MSE.It should be noted that the mean squared error values cannot be averaged to determine the RMSE. Beginners frequently make this mistake, which is an illustration of Jensen's inequality.The square root is, as you might remember, the opposite of the square operation. The square operation is used by MSE to eliminate the sign from each error number and to penalise significant errors. Although it assures that the outcome is still positive, the square root reverses this action.Using the mean squared error() function from the scikit-learn library, you may determine the root mean squared error between your predicted and expected values.By setting the "squared" argument to False, we may instruct the function to compute the square root of the MSE rather than the MSE itself, which is what it does by default.The function returns the mean squared error value from a one-dimensional array or list of expected values and projected values.

Conclusion

This study used machine learning techniques to forecast stock prices for a corporation with a higher degree of accuracy and reliability. The implementation of the innovative LSTM Model as a method of determining stock prices is the researchers' main contribution. Both approaches have improved prediction accuracy, resulting in successful outcomes, with the LSTM model proving to be more effective. The findings,

which show that machine learning techniques can be used to anticipate the stock market more effectively and accurately, are highly encouraging. By using a much larger dataset than the one being used right now, the stock market prediction system's accuracy can be significantly increased in the future. Additionally, other machine learning models that are currently under development should be examined to see how accurate they are. A potentially interesting area is sentiment analysis using machine learning to determine how news impacts a company's stock price. It is also possible to employ other deep learning-based models for prediction.

REFERENCES

[1] . Usmani, S. H. Adil, K. Raza and S. S. A. Ali, "Stock market prediction using machine learning techniques," 2016 3rd International Conference on Computer and Information Sciences (ICCOINS), Kuala Lumpur, 2016, pp. 322-327

[2] . 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] . L. Siew and M. J. Nordin, "Regression techniques for the predictio of stock price trend," 2012 International Conference on Statistics in Science, Business and Engineering (ICSSBE), Langkawi, 2012, pp.

[4] 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.

[5] . Liu, G. Liao and Y. Ding, "Stock transaction prediction and analysis based on LSTM," 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA), Wuhan, 2018,

[6] 787-2790.[8] 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

[7] . A. Althelaya, E. M. El-Alfy and S. Moha 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

[8] . A. Althelaya, E. M. El-Alfy and S. Moha 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

[9] abitza, F., Locoro, A., Banfi, G. (2018). Machine learning in orthopedics: a literature review. Frontiers in Bioengineering and Biotechnology, 6, 75.

[10] havan, P. S., Patil, S. T. (2013). Parameters for stock market prediction. International Journal of Computer Technology and Applications, 4(2), 337.

[11] hiu, D. Y., Chen, P. J. (2009). Dynamically exploring internal mechanism of stock market by fuzzy-based support vector machines with high dimension input space and genetic algorithm. Expert Systems with Applications, 36(2), 1240-1248.

[12] hong, E., Han, C., Park, F. C. (2017). Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. Expert Systems with Applications, 83, 187-205

[13] ai, W., Wu, J. Y., Lu, C. J. (2012). Combining nonlinear independent component analysis and neural network for the prediction of Asian stock market indexes. Expert systems with applications, 39(4), 4444-4452

[14] as, S. P., Padhy, S. (2012). Support vector machines for prediction of futures prices in Indian stock market. International Journal of Computer Applications, 41(3).

[15] ash, R., Dash, P. K. (2016). A hybrid stock trading framework integrating technical analysis with machine learning techniques. The Journal of Finance and Data Science, 2(1), 42-57.

[16] nke, D., Thawornwong, S. (2005). The use of data mining and neural networks for forecasting stock market returns. Expert Systems with applications, 29(4), 927- 940.

[17] oodhue, D. L., Thompson, R. L. (1995). Task-technology fit and individual performance. MIS quarterly, 213-236.

[18] uresen, E., Kayakutlu, G., Daim, T. U. (2011). Using artificial neural network models in stock market index prediction. Expert Systems with Applications, 38(8), 10389-10397

[19] olland, J. H. (1992). Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence. MIT press

[20] asic, T., Wood, D. (2004). The profitability of daily stock market indices trades based on neural network predictions: Case study for the SP 500, the DAX, the TOPIX and the FTSE in the period 1965–1999. Applied Financial Economics, 14(4), 285-297.

[21] im, H. J., Shin, K. S. (2007). A hybrid approach based on neural networks and genetic algorithms for detecting temporal patterns in stock markets. Applied Soft Computing, 7(2), 569-576.

[22] im, K. J., Han, I. (2000). Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index. Expert systems with Applications, 19(2), 125-132

[23] im, K. J., Lee, W. B. (2004). Stock market prediction using artificial neural networks with optimal feature transformation. Neural computing applications, 13(3), 255-260.

[24] im, M. J., Min, S. H., Han, I. (2006). An evolutionary approach to the combination of multiple classifiers to predict a stock price index. Expert Systems with Applications, 31(2), 241-247.

[25] umar, L., Pandey, A., Srivastava, S., Darbari, M. (2011). A hybrid machine learning system for stock market forecasting. Journal of International Technology and Information Management, 20(1), 3.

[26] ee, K. H., Jo, G. S. (1999). Expert system for predicting stock market timing using a candlestick chart. Expert systems with applications, 16(4), 357-364.

[27] ee, M. C. (2009). Using support vector machine with a hybrid feature selection method to the stock trend prediction. Expert Systems with Applications, 36(8), 10896-10904

[28] i, X., Xie, H., Wang, R., Cai, Y., Cao, J., Wang, F., ... Deng, X. (2016). Empirical analysis: stock market prediction via extreme learning machine. Neural Computing and Applications, 27(1), 67-78.

[29] iao, Z., Wang, J. (2010). Forecasting model of global stock index by stochastic time effective neural network. Expert Systems with Applications, 37(1), 834-841.

[30] alhotra, R. (2015). A systematic review of machine learning techniques for software fault prediction. Applied Soft Computing, 27, 504-518.

[31] osavi, A., Ozturk, P., Chau, K. W. (2018). Flood prediction using machine learning models: Literature review. Water, 10(11), 1536

[32] , J., Lee, J., Lee, J. W., Zhang, B-T. (2006). Adaptive stock trading with dynamic asset allocation using reinforcement learning. Information Sciences, 176(15), 2121-2147

[33] u, P., Wang, H. (2009). Prediction of stock market index movement by ten data mining techniques. Modern Applied Science, 3(12), 28-42.

[34] atel, J., Shah, S., Thakkar, P., Kotecha, K. (2015). Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. Expert Systems with Applications, 42(1), 259-268.

[35] ierdzioch, C., Risse, M. (2018). A machine-learning analysis of the rationality of aggregate stock market forecasts. International Journal of Finance Economics, 23(4), 642-654

[36] ound, J. (2019, December 24). Global stock markets gained 17 trillion in value in 2019. Retrieved from https://www.cnbc.com/2019/12/24/global-stock-marketsgained-17-trillion-in-value-in-2019.html.

[37] ian, B., Rasheed, K. (2007). Stock market prediction with multiple classifiers. Applied Intelligence, 26(1), 25-33

[38] chumaker, R. P., Chen, H. (2009). Textual analysis of stock market prediction using breaking financial news: The AZFin text system. ACM Transactions on Information Systems (TOIS), 27(2), 12.

[39] chumaker, R. P., Chen, H. (2010). A discrete stock price prediction engine based on financial news. Computer, 43(1), 51-56.

[40] chumaker, R. P., Chen, H. (2010). A discrete stock price prediction engine based on financial news. Computer, 43(1), 51-56.

, X., Xie, H., Wang, R., Cai, Y., Cao, J., Wang, F., ... Deng, X. (2016). Empirical analysis: stock market prediction via extreme learning machine. Neural Computing and Applications, 27(1), 67-78.on cites