International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 01 | January - 2024 SJIF Rating: 8.176 ISSN: 2582-3930

Prediction of Stock Prices Using Machine Learning

Adesh More
Department of computer Engineering, VPKBIET Baramati
moreadesh1960@gmail.com

Prathamesh Pawar
Department of computer Engineering, VPKBIET Baramati
prathameshpawar1877@gmail.com

Shoyeb Sayyed
Department of computer Engineering, VPKBIET Baramati
sayyadshoyeb37@gmail.com

Tejas Pagale
Department of computer Engineering, VPKBIET Baramati
Tejasp9339@gmail.com

Abstract: Researchers have been examining various approaches for making accurate stock market price predictions. Effective prediction systems give traders more insights into data, including future trends. Additionally, since the analysis provides future market conditions, investors stand to gain significantly. Forecasting with machine learning algorithms is one such technique. Improving the quality of output from stock market predictions using stock value is the aim of this project. Many researchers have created different strategies to address this issue; conventional methods, for example, employ artificial neural networks to uncover latent patterns and categorize data that is utilized for stock market forecasting. This project suggests an alternative approach to stock market price prediction.

INTRODUCTION

Stockbrokers, traders, and investors purchase and sell equities on the stock market. Purchasing a stock with a high possibility of price increase and selling it with a high probability of price decrease is the primary method employed by financial experts. In general, there are two methods for predicting the stock market. Among these is fundamental analysis, which is predicated on the strategy of an organization and basic information like as expenses, annual growth rates, and market position. The second approach, referred to as technical analysis, is centered on previous stock prices and values. Financial analysts had to constantly forecast stock markets in the past in order to predict future stock values. In order to increase their investment returns, investors are experimenting with various methods to learn more about various companies. The stock market is regarded as one of the most profitable financial venues since it is used to trade large amounts of capital on a regular basis. Many people make large financial investments in the stock market, but occasionally they lose a great deal of money because they rely on stockbrokers, who give investors advice based on time series, technical, and fundamental analysis. Investors have been searching for a clever solution to these issues. Because it is so important to predict stock prices, this is where stock price prediction comes into play.

The basic goal of stock price prediction is to forecast future financial results with accuracy. Many traders are implementing machine learning algorithms into their respective professions as they have shown promise in a number of businesses in recent years. ML has the potential to change the game.

This study conducts experiments using several machine learning (ML) methods to forecast a stock's closing price. LSTM, SVM, LR, and Multi-layer Preceptron are the machine learning algorithms that we will use.

RELATED WORK

In paper [1], The researcher suggests a method for predicting stock prices using machine learning algorithms. The study highlights how difficult it is to predict stock prices because of the market's volatility and the wide range of variables that influence them. The financial market generates a lot of data, which makes it challenging for traders to identify trends and come up with the best strategies for making decisions, according to the author. Two methods are used by the proposed system to operate: regression and classification. Classification predicts whether the stock price will rise or fall the following day, whereas regression predicts the closing price of a stock. The researcher notes that the system predicts stock prices based on a number of variables, including the opening price, closing price, highest price, lowest price, and volume of shares traded. The fact that Yahoo Finance provides the system's data is another point made by the author. Since the data is in an unprocessed state, analysis is not possible. As a result, the researcher notes that any missing values and outliers are removed from the data through preprocessing. The system predicts stock prices using a variety of machine learning algorithms, including Support Vector Machine, Random Forest, and Linear Regression.

In paper [2], The use of machine learning algorithms—more especially, SVM and linear regression—to forecast stock market trends is covered in this paper. The primary aim is to use historical data analysis to lower the chances of failures in the future. The significance of accurately labeled and updated datasets for precise predictions is emphasized in the paper. Additionally, it evaluates the effectiveness of SVM and Linear Regression in stock market analysis, concluding that the latter is more appropriate for this kind of work. In summary, the study highlights how machine learning can be used to forecast stock prices and how this can affect trading and financial services.

In paper [3], In order to predict stock prices, the researchers compared machine learning techniques with conventional statistical methods. The study emphasizes how difficult it is to predict stock prices because they are not fixed in time. Thus, the goal of the research is to determine the most effective technique for stock closing price prediction. Numerous techniques are used in the study, including the following: Random Forest, Support Vector Machine, Lasso, Ridge, K-Nearest Neighbors, Exponential Smoothing, Simple Moving Average, Weighted Moving Average, Naive Approach, Linear Regression, Single Layer Perceptron, Multi-layer Perceptron, and Long Short Term Memory. Using these techniques, the researchers made predictions, then evaluated the accuracy and performance of their predictions. Upon analyzing each technique separately, it was discovered that the machine learning strategy—particularly the neural network models—was the most accurate for predicting stock prices. The study also shows that in terms of prediction accuracy and performance, the machine learning approach performs better than the conventional statistical approaches.

In paper [4], The researchers' suggestion is to forecast stock prices using an LSTM neural network. The prediction effect of the model on stock time series was verified through experiments with both single and multi-feature input variables. The outcomes demonstrated the high prediction accuracy of the LSTM model for stock prices. The performance of the univariate feature input model and the multivariate feature input model was

compared, and the researchers also examined the prediction results. The multivariate model proved to be more effective, as evidenced by its faster convergence and smaller loss function value. This implies that analyzing and forecasting volatile and unstable stock prices benefits from the use of multiple features in the model training process. The researchers also talked about how crucial data normalization is to the analysis of stock data. They emphasized that stock prices and volume can vary greatly in magnitude and that results of directly analyzing the original data can be skewed. They underlined the necessity of normalizing the data in order to remove the influence of data magnitude on the analysis's conclusions.

In paper [5], The researchers suggest a model to enhance trading signals in the stock market by combining machine learning and moving averages. Moving averages, according to the researchers, are a helpful tool for spotting trends, but they are also lagging indicators, which means that they provide trading signals only after a significant change in price has occurred. The researchers suggest applying machine learning techniques to anticipate the moving average crossover ahead of time in order to get around this drawback. This will shorten the lag and increase the precision of trading signals.

In paper [6], In order to predict stock price trends, the researchers developed a novel CNN methodology that can be used to simulate primary image output and incorporate data straight into an image. The proposed model can train the network weights using image chunks that are provided as input. As a result, only features that are helpful in helping to manage the identity of extreme values (high/low) in the market through signal validation will be extracted. The input is historical data obtained directly from the Taiwan Stock Market (TSM).

In paper [7], In terms of accuracy in predicting the behavior of the stock market, the researchers have compared the stacked LSTM network model to other time series analysis models. Based on the obtained results, the model was able to predict future stock market behavior with an accuracy of 53.6%. Additionally, using historical data from the Pakistani company OGDCL, the study compares the use of ARIMA for stock price prediction. With an error estimate of 0.63 in monetary units, the results demonstrated the effectiveness of ARIMA for short-term forecasting. In summary, the study offers significant understanding of how to apply a stacked LSTM network model to forecast stock market behavior and contrasts it with alternative time series analysis models.

In paper [8], The difficulties in forecasting stock prices and the ways in which deep learning methods like LSTM can increase precision have been discussed by the researchers. Traditional methods like support vector regression and linear regression are mentioned in the paper as not being ideal for the investment goal. This paper suggests an approach that first employs LSTM to refine sequential embedding using historical data, followed by Graph Convolution. The analysis of the experiment indicates that this approach leads to a significant improvement in return ratio.

STOCK PREDICTION TECHNIQUES

The algorithms generally used for stock price prediction are as follows:

1. Simple Linear Regression
2. Support Vector Machine (SVM)
3. Multi-Layer Perceptron
4. Long Sort Term Memory (LSTM)

1. Simple Linear Regression:

Simple linear regression is a statistical method used to model the relationship between two quantitative variables: one independent variable (often denoted as X) and one dependent variable (often denoted as Y). It assumes a linear relationship between the variables, meaning that changes in the independent variable are associated with a constant change in the dependent variable.

The model for simple linear regression can be expressed as:

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

Here, Y is the dependent variable.

X is the independent variable.

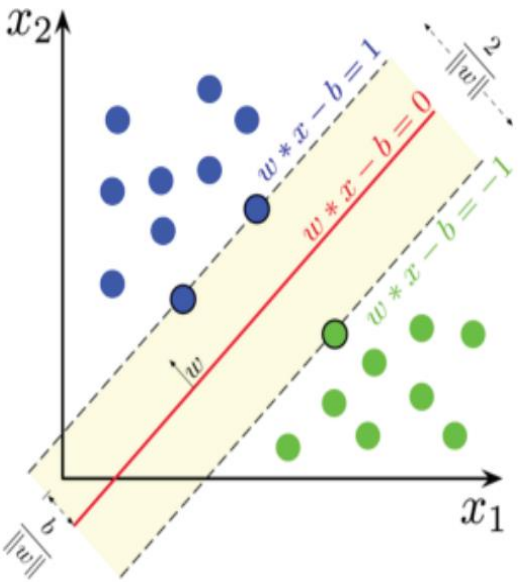
B_0 is the y-intercept.

B_1 is the slope of the regression line.

Ɛ is the slope of the regression line.

2. Support Vector Machine (SVM):

SVM is a machine learning algorithm that is used in solving some of the regression and prediction problems. It is primarily based on statistical learning approach. In SVM we to predict results we find out the place between which the value might range for the input. In SVM we analyze the two classes for dividing them with a hyperplane, it is a line that splits the data of the two classes such that value between the last points of the classes is maximum it is called maximum margin. Now using the hyperplane and the support vectors we achieve the widest street, which is in further enhance towards SVM approach.



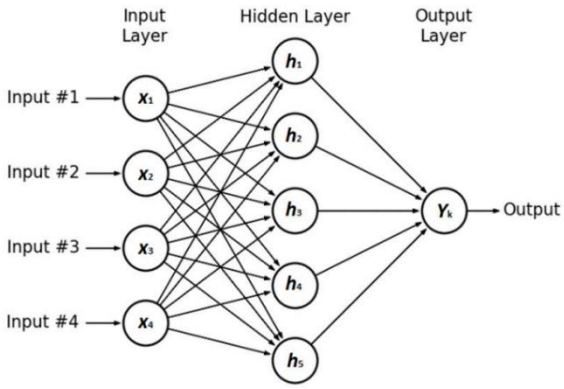
Here, w = vectors

x = variables

b = biased term (wo)

3. Multi-Layer Perceptron:

Multi-Layer Preceptron is a type of artificial neural network that consists of multiple layers of interconnected nodes. It is a feedforward neural network, meaning that information flows in one direction from the nodes, through the hidden layers, to the output nodes.



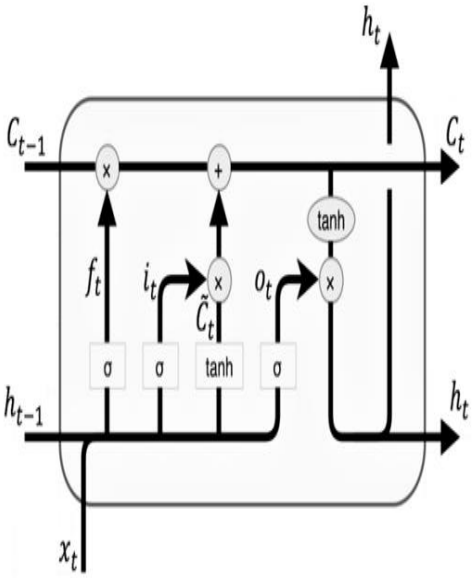
Output calculation formula for output layer is as follows:

Here, Y_k = Final Output

Θ_k = Threshold

Φ = Activation function

n = Number of neurons



4. Long Short Term Memory:

Long Short-Term Memory (LSTM) is a type of recurrent neural network (RNN) architecture designed to capture long-term dependencies in sequential data. It was introduced to address the limitations of traditional RNNs, which struggle with learning dependencies over long sequences. LSTMs are a powerful tool for

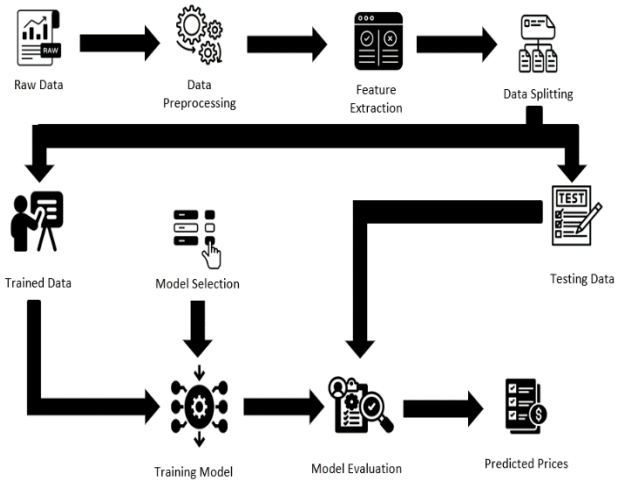
$$o_t = \sigma(W_O[h_{t-1}, x_t] + b_o)$$
$$t_t = o_t * \tanh(c_t)$$

modeling sequences, and their ability to capture long-term dependencies makes them a popular choice for a wide range of tasks involving sequential data. They have been foundational in the development of many state-of-the-art models in fields like natural language processing and time series analysis.

Here,

$$f_t = \sigma(W_f \bullet [h_{t-1}, x_t] + b_f)$$
$$i_t = \sigma(W_i \bullet [h_{t-1}, x_t] + b_i)$$
$$\tilde{c}_t = \tanh(W_C \bullet [h_{t-1}, x_t] + b_c)$$
$$c_t = f_t * c_{t-1} + i_t * \tilde{c}_t$$

SYSTEM ARCHITECTURE



COMPARITIVE STUDY OF EXISTING APPROACHES

Model	Pros	Cons	Suitable for
Linear Regression	Simple and easy to interpret.	Assumes linear relationship.	Short-term trends, basic analysis.
SVM	Effective in high-dimensional spaces.	Can be computationally expensive.	Non-linear data, classification.
Multi-Layer Perceptron	Can capture complex non-linear relationships.	Sensitive to hyperparameters, may overfit.	Complex relationships, general ML.
LSTM	Captures long-term dependencies.	Can be computationally intensive to train.	Time series with long-term patterns.

CONCLUSION

In this review paper, we have delved into the application of machine learning algorithms in the domain of stock market prediction, focusing on Linear Regression, Support Vector Machines (SVM), Multi-Layer Perceptrons (MLP), and Long Short-Term Memory (LSTM) networks. Through a comprehensive analysis of these methodologies, we have gained valuable insights into their strengths, limitations, and suitability for different aspects of stock market forecasting.

REFERENCE

[1] Srinath Ravikumar, Prasad Saraf, "Prediction of Stock Prices using Machine Learning (Regression,Classification) Algorithms" , 2020.

[2] Kunal Pahwa, Neha Agarwal, "Stock Market Analysis using Supervised Machine Learning", 2019.

[3] Indronil Bhattacharjee, Pryonti Bhattacharja, "Stock Price Prediction: A Comparative Study between Traditional Statistical Approach and Machine Learning Approach", Jun 5-7, 2019.

[4] Jingyi Du, Qingli Liu , Kang Chen, Jiacheng Wang, "Forecasting stock prices in two ways based on LSTM neural netwrok" , ITNEC 2019.

[5] Shoban Dinesh, Nithin Rao R, S P Anusha, "Prediction of Trends in Stock Market using Moving Averages and Machine Learning.", Apr, 2021.

[6] Jimmy Ming-Tai Wu, Zhongcui Li, Gautam Srivastava, "A CNN-based Stock Price Trend Prediction with Futures and Historical Price", 2020.

[7] Gourav Bathla, "Stock Price prediction using LSTM and SVR", 2020.

[8] Samuel Olusegun Ojo, Pius Adewale Owolawi, Maredi Mphahlele, "Stock Market Behaviour Prediction using Stacked LSTM Networks", 2020.