

A TECHNICAL SURVEY ON FORECASTING STOCK PRICE USING MACHINE LEARNING

P. Sobana¹, Mr. S Prabhu², Dr. S.Karuppusamy³

¹PG Scholar, ²Associate Professor, ³Professor

Department of Computer Science and Engineering,
Nandha Engineering College, Erode, Tamandu

ABSTRACT

Stock market could be a complicated and difficult system wherever folks can either gain cash or lose their entire life savings. The aim in stock Market Prediction is to predict the future price of the money stocks of a corporation. The latest trend in inventory marketplace prediction technology is that using machine learning that makes predictions supported the values of modern stock marketplace indices via training on their previous values. Machine learning itself employs completely different models to form prediction easier and authentic. Varied approaches and also the results of past years are compared based on methodologies, datasets and potency so it's delineated within the type of a table. The paper focuses on the employment of Regression primarily based Machine learning to predict stock values. Factors thought-about are open, close, low and high.

Keywords: Machine Learning, Regression algorithms, Python programming

1. INTRODUCTION

Stock Market analysis has been an important space of analysis and is one in all the highest applications of machine learning. Prediction and analysis of the stock market is one of the foremost sophisticated tasks to try to do. There are numerous reasons for this, like the market volatility then several different dependent and independent factors for deciding the worth of a specific stock within the market. However, with the arrival of Machine Learning and its strong algorithms, the newest market research and stock market Prediction traits have begun out incorporating such strategies in understanding the stock market information. In short, Machine Learning Algorithms are getting used extensively by means of several organizations in analyzing and predicting stock values. This paper will undergo the usage of Multiple Regression algorithm in forecasting the Stock Price Prediction via way of means of thinking about the elements open, close, excessive and low.

1.1 Prediction:

The attempt of trying to decide the destiny price of the inventory marketplace is referred to as inventory marketplace Prediction. Predicting stock market expenses is a complicated task as it includes determination of future fee of stock unit traded on an exchange. The successful prediction of a stock's unborn price could grant significant profit. Due to the considerable quantity of cash worried and quantity of proceedings that take area each shake, there have to be rigor among the volumes of predictions made. The technical analysis approach recommends generating predictions based on the historical price values of opted stocks. The data set of the stock market prediction model contains details like the closing price, opening price, the data and various other parameters

that are took to prophesy the end variable which is the cost in a given day. Stock Market prices can be prophesied based on two ways: Current prices of the stocks and both the prices and the news rubrics.

- (i) **Current prices of the stocks** – Mostly prices change from day to day on a fixed amount or at a constant rate. These are the kind of common mutual funds where quantity if invested, will be compounded manually. This is not of definite interest as there is zero use of a machine to guess the coming price. Only a calculator is enough.
- (ii) **Both the prices and the news captions** – The prices dominated to these will vary from time to time as grounded on their bearing. Say a company fired a product which attain the market and got really connected to people. Obviously, sales will be amplifying for that company and the investor who invested in that individual company will be gainful. For these kinds of calculi, we require some tool to be effective. These predictions are played using several conservative forms such as Traditional Time Series, Technical Analysis Methods, Machine Learning Methods, and Fundamental Analysis Methods. The choice of the forenamed methods is grounded on the type of tool applied and the data upon which the tool is administered.

1.2 Assumption of Regression Model:

- **Linearity:** The relationship between dependent and independent variables ought to be linear.
- **Homoscedasticity:** Constant variance of the mistakes has to be maintained.
- **Multivariate normality:** Multiple Regressions assumes that the residuals are typically distributed.
- **Lack of Multicollinearity:** It is assumed that there is very little or no multicollinearity in the data.

2. LITERATURE SURVEY

Mojtaba Nabipour, Pooyan Nayyeri, Hamed Jabani, Shahab S, and Amir Mosavi carried out a survey of evaluating nine machine learning models (Decision Tree, Random Forest, Adaptive Boosting (Adaboost), eXtreme Gradient Boosting (XGBoost), Support Vector Classifier (SVC), Naïve Bayes, K-Nearest Neighbors (KNN), Logistic Regression and Artificial Neural Network (ANN)) and two effective deep learning methods (Recurrent Neural Network (RNN) and Long short-term memory (LSTM)). Ten technical indicators from ten years of historical data are the input values, and two ways are meant for using them. Each prediction model is evaluated through three metrics supported the input ways in which. The analysis result suggests that for the ceaseless data, RNN and LSTM outperform other auguring models with a great difference.

Jai Prakash Verma and all presents an article learns the model from Indian National Stock Exchange (NSE) data obtained from Yahoo API to forecast stock prices and goals to make a profit over time. In this paper, two discrete algorithms and methodologies are analyzed to prognosticate stock market trends and iteratively upgrade the model to gain better preciseness. Results are proving that the proposed pattern-based customized algorithm is greater accurate (10 to 15%) in comparison to other two machine learning techniques, which are also hyped as the time window increases.

Tejas Mankar, Tushar Hotchandani, Manish Madhwani, Akshay Chidrawar, Lifna C.S addressed 3 challenges (1) Historical Twitter data can't be obtained, unless it's far stored with the aid of using someone, thus records must be gathered over a period of a hard and fast wide variety of months beginning from the prevailing date and time; (2) it's important to clear out needed data from the stream of unrelated tweets; and (3) Authentication is needed for gaining access to real time Twitter data. They completed a comparative and concluded that Support Vector Machine proven to be the foremost economical and possible model in predicting the stock worth movement, in favor of the sentiments of the tweets.

Bikrant Bikram P. Maurya, Ayush Ray, Aman Upadhyay, provided a paper that explains the prediction of an inventory using Machine Learning. The input parameters embody -open, high, low, close rate, buying and selling volume, Price to Earning Ratio, MA, MACD for greater accuracy. The Machine Learning algorithm, Random Forest Regression has been brought in Python programming language that is utilized to augur the exchange. The algorithm has been used at the historic stock information in conjunction with web- scraping approach that has been implemented to seize contemporary marketplace information of the stock. The recursive training model takes its foreseen price as input to predict additional additional long-time future stock rates.

Minal Khan furnished a machine-based technique changing the culture but unreliable time collection forecasting to expect the shares from the beyond historic facts of a massive organization and decide the underlying styles to enhance upon the same. After an extended deliberation, multiple traditional gadgets studying algorithms has been used that best predicts the inventory fees and as a result assist in figuring out the underlying styles to attract end from to plan techniques in enhancing the shares.

Sathyabama S, Stemina S C, SumithraDevi T and Yasini N used strategies together with Random Forest and NB (Naive Bayes) at the dataset. And those algorithms have given output one at a time primarily based totally on its performance. Naïve Bayes classifier has created more accuracy while in comparison with Random Forest. Use of lately delivered machine learning strategies inside the prediction of shares have yielded promising outcomes and thereby marked the usage of them in worthwhile alternate schemes. It has brought about the belief that it's far feasible to forecast the worth of the inventory markets with additional accuracy and performance with the usage of machine learning techniques.

Irfan Ramzan Parray, and all used 3 Machine Learning algorithms employed on this paper are Support vector Machine, perceptron, and logistic regression, for predicting the succeeding day trend of the shares. For the experiment, dataset from concerning fifty shares of Indian National Stock Exchange's NIFTY 50 index turned into taken, with the aid of using amassing inventory facts from January 1, 2013, to December 31, 2018, and ultimately with the aid of using the calculation of a few technical indicators. It is stated that the common accuracy for the prediction of the trend of 50 shares received with the aid of using Support vector Machine is 87.35%, perceptron is 75.88%, and logistic regression is 86.98%. Since the inventory facts are time series data, every other dataset is ready by reorganizing preceding dataset into the supervised learning layout that improves the accuracy of the prediction system which stated the outcomes with Support vector Machine of 89.93%, perceptron of 76.68%, and logistic regression of 89.93, respectively.

Jingyi Shen and M. Omair Shafiq performed complete reviews on often used machine learning models and that our projected answer outperforms because of the complete function engineering that we tend to designed. The system achieves basic excessive accuracy for inventory marketplace trend prediction. With the targeted layout and assessment of prediction time period lengths, feature engineering and statistics pre-processing methods, the work contribute to the inventory evaluation research community each withinside the economic and technical domains.

Mariam Moukalled Wassim El-Hajj Mohamad Jaber proposed an automatic trading system that integrates mathematical functions, machine learning, and different outside elements consisting of news' sentiments for the reason of reaching higher inventory prediction accuracy and issuing worthwhile trades. Significantly, the purpose is to decide the rate or the trend of a certain inventory for the coming end-of-day thinking about the primary numerous trading hours of the day. To reap this goal, they skilled conventional machine learning algorithms and created/skilled more than one deep learning models taking into account the significance of the applicable news.

Wasiat Khan, Mustansar Ali Ghazanfar, Muhammad Awais Azam, Amin Karami, Khaled H. Alyoubi, Ahmed S. Alfakeeh use algorithms on social media and economic news data to find out the effect of this statistics on inventory marketplace prediction accuracy for ten next days. For enhancing overall performance and quality of predictions, feature selection and spam tweets reduction are done at the data sets. Moreover, we carry out experiments to locate such inventory markets which are hard to predict and people which are extra stimulated via way of means of social media and economic information. We evaluate consequences of various algorithms to discover a steady classifier. Finally, for reaching most prediction accuracy, deep learning is used and a few classifiers are ensembled.

Prof. Ashwini Kanade, Sakshi Singh, Shweta Rajoria³, Pooja Veer, Nayan Wandile, Prof. Ashwini Kanade, Sakshi Singh, Shweta Rajoria, Pooja Veer, Nayan Wandile suggest a framework using Long Short-Term Memory machine learning algorithm and adaptive inventory technical signs for efficient forecasting through the use of diverse parameters received from the historic statistics set taken into consideration for a selected company. This set of rules works on historic statistics retrieved from Yahoo Finance. For prediction of percentage fee using Long Short-Term Memory, there are modules, one is training session and other is predicting charge primarily based totally on formerly skilled statistics.

Edgar P. Torres P, Myriam Hernández-Álvarez, Edgar A. Torres Hernández, and Sang Guun Yoo experimented with inventory marketplace information of the Apple Inc. using random trees and multilayer perceptron algorithms to carry out the predictions of last costs. An accuracy evaluation becomes additionally carried out to decide how beneficial can those kinds of supervised machine learning algorithms may be withinside the economic field. These kinds of research can also be researched with information from the Ecuadorian inventory marketplace exchanges i.e., Bolsa de Valores de Quito (BVQ) and Bolsa de Valores de Guayaquil (BVG) to assess the effectiveness of the algorithms in much fewer liquid markets and probably assist lessen inefficiency charges for marketplace individuals and stakeholders.

M Umer Ghani, M Awais and Muhammad Muzammul used Machine Learning Algorithm mainly attention on Linear Regression (LR), Three month Moving Average(3MMA), Exponential Smoothing (ES) and Time Series Forecasting using MS Excel as high-quality statistical device for graph and tabular illustration of prediction outcomes. We tend to obtain knowledge from Yahoo Finance for Amazon (AMZN) inventory, AAPL inventory and GOOGLE inventory after implementation LR we efficaciously expected inventory marketplace trend marketplace trend for subsequent month and additionally measured accuracy in step with measurements. Shubham Jain, Mark Kain supplied a paper on the way to expect inventory values supported the info of NY Times of 10 years using Machine Learning algorithms: Logistic Regression, Random Forest and Multilayer Perceptron (MLP). They finished that MLP is higher than the alternative algorithms because, inside a positive range, the distinction rate and expected rate is pretty small compared to the ones in Logistic Regression and Random Forest.

Subhadra Kompella, Kalyana Chakravarthy Chilukuri expected inventory marketplace which takes input as the rate of inventory and information heading. Sentiment analysis is employed to calculate the polarity rating after which use it similarly in detecting the sort of article has a fantastic or terrible effect toward the inventory and people may be used similarly withinside the evaluation. The acquired ratings are used to calculate the costs of stock and to finish the ones inputs as a fixed we used the exponential moving average approach and stored because the effect of stock is efficaciously determined. The information after calculating is updated and exhibited to the consumer as a graph. Finally, Random Forest Algorithm is implemented and as in comparison with logistic regression for performance.

M. Ananthi, K. Vijayakumar projected a system that predicts inventory rate of any enterprise cited with the aid of using the for the following few days. The proposed system generates alerts at the candlestick graph that permits to predict marketplace movement to an adequate stage of accuracy in order that the consumer is in a position to evaluate whether or not a stock may be a 'Buy/Sell' and whether or not to brief the stock or pass lengthly by delivery. The prediction accuracy of the stock market has analyzed and advanced to 85% the use of machine learning algorithms. MIN WEN, PING LI, LINGFEI ZHANG, AND YAN CHEN added a brand-new approach to simplify noisy-filled financial temporal collection through series reconstruction by investing motifs (frequent patterns), after which make use of a convolutional neural network to seize spatial shape of time series. The experimental outcomes display performance of our proposed approach in feature engineering and outperformance with 4percent accuracy development as in comparison with the conventional sign procedure strategies and frequency commerce patterns modeling method with deep learning in inventory trend prediction.

Nirbhey Singh Pahwa, Nirbhey Singh Pahwa, Vidhi Soni, Deepali Vora recommends use of linear regression and logistic regression for inventory prediction and inventory evaluation and this have a look at SVM to acquire correct outcomes. A constraint to this end is the need of the dataset utilized in prediction to be classification friendly.

Srinath Ravikumar, Prasad Saraf proposed a system that works in two strategies – Regression and Classification. In regression, the system predicts the last rate of inventory of an enterprise, and in classification, the system predicts whether or not the last rate of inventory will boom or lower the following day. The input dataset acquired from Yahoo Finance become correctly preprocessed and diverse huge attributes have been brought to the dataset like momentum, volatility and sector details.

3. COMPARITIVE ANALYSIS

S.No	Topic	Methodology/ Techniques	Parameter	Future work
------	-------	-------------------------	-----------	-------------

1.	Predicting Stock Market Trends Using Machine Learning and Deep Learning Algorithms Via Continuous and Binary Data; a Comparative Analysis	(Decision Tree, Random Forest, Adaboost, XGBoost, SVC, Naïve Bayes, KNN, Logistic Regression and ANN) and two deep learning algorithms (RNN and LSTM).	Better performance Predicts task of stock market	To conduct valuable researches in stock market.
2.	Evaluation of Pattern Based Customized Approach for Stock Market Trend Prediction with Big Data and Machine Learning Techniques	SMP model, Support Vector Machine (SVM), Artificial Neural Network (ANN)	Achieves maximum accuracy up to 10-15%	Considering the heterogeneity of data, the problem would be considered under big data problem.
3.	Stock Market Prediction based on Social Sentiments using Machine Learning	Naïve Bayes, SVM, Sentiment Analysis	Predicts efficient stock price movement	To study and incorporates the other factors that affect public mood.
4.	Recursive Stock Price Prediction with Machine Learning and Web Scrapping for Specified Time Period	Web scraping technique, Random Forest, Regression algorithm	Increase in future prediction with decrease in predicted stock price. (Accuracy improved)	To tackle real world and real time stock problems more efficiently.
5.	A Review Paper on Stock Prediction Using Machine Learning Algorithms	Regression, Classification, Multi-Layer Perceptron model, Support Vector Machine model, and Partial Least Square Classifier	Predicts close price Predicts close price whether close price will increase or decrease the next day.	to train the algorithm on huge collection of data.
6.	Intelligent Monitoring and Forecasting Using Machine Learning Techniques	Navies Bayes Classifier, Random Forecast algorithm	Gives maximum accuracy when compared with other random forecast algorithms. The problem of overfitting is ignored.	Stock Market price forecasting system can be enhanced by using a bigger dataset in future.
7.	Time series data analysis of stock price movement using machine learning techniques	Logistic Regression, SVM, Perceptron Model of ANN	2% increase in accuracy for testing training split method.	The work can be expanded by adding noise more historical data and calculation of more technical indicators as features.
8.	Short term stock market price trend prediction using a comprehensive deep learning system	Recursive feature Elimination (RFE), Principal Component Analysis, LSTM	Leverages dimensionality Reduction	In-depth research into the technical indices which influence the irregular term lengths.
9.	Automated Stock Price Prediction Using	Recurrent neural network (RNN), feed forward neural network (FFNN), support vector machines (SVM) and	Predicts the direction of today's close price with respect to yesterday's close price. Improves	Enhance the prediction of the exact price.

	Machine Learning	support vector regression (SVR)	learning rate, learning process.	
10.	Stock market prediction using machine learning classifiers and social media, news	Feature selection module and PCA class from the decomposition module, SelectKBest	Dimensionality reduction or feature selection is performed.	Use more systematic technique for determining stock relevant keyword for searching social media and news.
11.	Machine Learning Model for Stock Market Prediction	Optimization algorithm, back propagation algorithm, LSTM, Sentiment Analysis	Calculates positive or negative influences of stock.	Research to predict the stock prices more accurately with wide range of inputs.
12.	Stock Market Data Prediction Using Machine Learning Techniques	Random tree, Multilayer perceptron algorithm	fits the actual historical price data very closely, errors are tolerable	To facilitate the ML algorithms and AI to Ecuadorian financial securities to further research the applications in other markets.
13.	Stock Market Prediction Using Machine Learning (ML) Algorithms	Linear Regression, Exponential Smoothing (ES), Time series forecasting	Improves accuracy	Further consideration of using many ML algorithms to predict stock price.
14.	Prediction for Stock Marketing Using Machine Learning	Linear Regression, Multilayer Perceptron Algorithm, Logistic Regression, Random Forest	Curves for average predicted price and predicted price becomes closer.	To use MLP further for improving accuracy.
15.	Stock market prediction using Machine learning methods	Random forest, Prediction, Time Series Analysis, Sentiment Analysis.	Achieves better score of regression metrics	To use Sentiment Analysis to detect the type of article has a positive or negative impact towards the stock.
16.	Stock market analysis using candlestick regression and market trend prediction (CKRM)	Regression, Classification, KNN regression algorithm, Candlestick pattern	Formulates the training sets, Improves Accuracy range.	A full-fledged Sentiment Analysis is to be implemented which helps to achieve greater performance in terms of market trend prediction.
17.	Stock Market Trend Prediction Using High-Order Information of Time Series	Motif Extraction, Conventional Neural Networks	Ups and downs prediction, More efficient in terms of computational complexity.	Boost the research on macroscopic pattern discovery in financial time series.
18.	Prediction of Trends in Stock Market using Moving Averages and Machine Learning	Regression algorithm	Reduces latency of the trading Improves accuracy	Further research to improve accuracy
19.	Stock Prediction using Machine Learning a Review Paper	Support Vector Regression (SVR), Linear regression, Logistic regression	Obtains accurate results	Further research to use more efficient tools to extract knowledge from data and perform prediction.

20.	Prediction of Stock Prices using Machine Learning (Regression, Classification) Algorithms	Regression, classification	Predicts close price Predicts close price whether close price will increase or decrease the next day.	The methods mentioned can be applied to real time data to get real time predictions of the stock value.
-----	---	----------------------------	--	---

4. CONCLUSION

Machine learning plays an important role in predicting the longer-term value of an organization. Stock exchange is the mitigation of risk through the spreading investments across multiple entities, that is achieved by the pooling of variety of little investments into an oversized bucket. This paper reviewed heaps of machine learning algorithms that suites most to predict the stock value additional accurately. Major papers reviewed regarding prediction of stock exchange trends, time series analysis, Short term stock market value prediction, usage of Candlestick Regression in stock value prediction. Considering this survey, the planned work is going to be associated with forecast the stock value exploitation Multiple Linear Regression.

5. REFERENCES

- [1]. Sameer Yadav 2017 Stock Market Volatility – A Study of Indian Stock Market. 2017 GJRA - Global Journal for Research Analysis.
- [2]. Mankar, T., Hotchandani, T., Madhwani, M., Chidrawar, A., & Lifna, C (2018). Stock Market Prediction based on Social Sentiments using Machine Learning. 2018 International Conference on Smart City and Emerging Technology (ICSCET). doi:10.1109/icscet.2018.8537242
- [3]. Maurya, B. B. P., Ray, A., Upadhyay, A Gour., B., & Khan, A. U. (2019). Recursive Stock Price Prediction with Machine Learning and Web Scrapping for Specified Time Period. 2019 Sixteenth International Conference on Wireless and Optical Communication Networks (WOCN). doi:10.1109/wocn45266.2019.8995080
- [4]. Minal Khan., Vijay Kumar Trivedi., Dr. Bhupesh Gour., (2020). A Review Paper on Stock Prediction Using Machine Learning Algorithms. 2020 The International journal of analytical and experimental modal analysis.
- [5]. Sathyabama S., Stemina S c., SumithraDevi T and Yasini N., (2021). Intelligent Monitoring and Forecasting using Machine Learning Techniques.
- [6]. Parray, I. R., Khurana, S. S., Kumar, M., Altalbe, A. A (2020). Time series data analysis of stock price movement using machine learning techniques. doi:10.1007/s00500-020-04957-x
- [7]. Shen, J., & Shafiq, M. O. (2020). Short-term stock market price trend prediction using a comprehensive deep learning system. Journal of Big Data, 7(1). doi:10.1186/s40537-020-00333-6
- [8]. M Umer Ghania., M Awaisa and Muhammad Muzammul., (2019). Stock Market Prediction Using Machine Learning (ML)Algorithms. Advances in Distributed Computing and Artificial Intelligence Journal 2019.
- [9]. Edgar P. Torres P., Myriam Hernández-Álvarez., Edgar A. Torres Hernández., and Sang Guun Yoo (2019). Stock Market Data Prediction Using Machine Learning Techniques.
- [10]. Ashwini Kanade., Sakshi Singh., Shweta Rajoria., Pooja Veer., Nayan Wandile (2020). Machine Learning Model for Stock Market Prediction. 2020 International Journal for Research in Applied Science & Engineering Technology (IJRASET).
- [11]. Wasiat Khan., Mustansar Ali Ghazanfar., Muhammad Awais Azam., Amin Karami., Khaled H. Alyoubi., Ahmed S. Alfakeeh (2020). Stock market prediction using machine learning classifiers and social media, news. 2020 Journal of Ambient Intelligence and Humanized Computing doi:10.1007/s12652-020-01839-w
- [12]. Mariam Moukalled., Wassim El-Hajj., Mohamad Jaber. Automated Stock Price Prediction Using Machine Learning.
- [13]. Mojtaba Nabipour., Pooyan Nayyeri., Hamed Jabani., Shahab S., Amir Mosavi (2020). Predicting Stock Market Trends Using Machine Learning and Deep Learning Algorithms Via Continuous and Binary Data; a Comparative Analysis IEEE.
- [14]. Jai Prakash Verma., Sudeep Tanwar., Sanjay Garg., Ishit Gandhi., Nikita H. Bachani (2019). Evaluation of Pattern Based Customized Approach for Stock Market Trend Prediction with Big Data and Machine Learning Techniques. 2019 International Journal of Business Analytics.
- [15]. Nirbhey Singh Pahwa., Neeha Khalfay., Vidhi Soni., Deepali Vora (2017). Stock Prediction using Machine Learning a Review Paper. 2017 International Journal of Computer Applications.
- [16]. Makridakis S, Spiliotis E, Assimakopoulos V (2018). Statistical and Machine Learning forecasting methods: Concerns and ways forward.

- [17]. Min wen., Ping li., Lingfei Zhang., Yan Chen (2019). Stock Market Trend Prediction Using High-Order Information of Time Series.
- [18]. M. Ananthi., K. Vijayakumar (2020). Stock market analysis using candlestick regression and market trend prediction (CKRM). 2020 Journal of Ambient Intelligence and Humanized Computing doi:10.1007/s12652-020-01892-5
- [19]. Subhadra Kompella., Kalyana Chakravarthy Chilukuri (2019). Stock Market Prediction Using Machine Learning Methods. 2019 International Journal of Computer Engineering & Technology (IJCET).
- [20]. Shubham Jain., Mark Kain Prediction for Stock Marketing Using Machine Learning. International Journal on Recent and Innovation Trends in Computing and Communication.

