



# Survey Paper On Stock Prediction Using Machine Learning Algorithms

<sup>1</sup>Lekshmi S, <sup>2</sup>Deepa Rajan, <sup>3</sup>Sam G Benjamin

<sup>1</sup>Student, <sup>2</sup>Assistant Professor, <sup>3</sup>Assistant Professor

<sup>1</sup>APJ Abdul Kalam Technological University,

<sup>1</sup>BMCE Sasthamcotta, Kollam, India

**Abstract:** Stock market is widely used in investment scheme promising high return but it has some risk. Stock returns are very fluctuating in nature. They depend upon various factors like previous stock prices, current market trends, financial news, social media etc. Many practices like technical analysis, fundamental analysis, time series analysis and statistical analysis etc. are used to predict the stock value but none of these procedures are proved as an allowable prediction tool. This paper surveys various machine learning techniques and algorithms employed to uplift the accuracy of stock price prediction.

**Index Terms - Machine Learning, Prediction, Support Vector Machine, Accuracy, Linear Regression.**

## I. INTRODUCTION

Stock market plays a vital role in the economic performance of a particular nation. Challenging risk in stock price prediction is its randomness in nature. Stock price prediction is one of the leading important issues faced by many stockholders. This increases the instability of funding our earned money with the fright of losing it. Hence, many are afraid to postulate within the stock exchange. Funding within the Stock markets generates remarkable revenue for the country. If a successful model for stock prediction is developed, we will gain intuition into the market movement over time, detecting the trends that might otherwise not have been perceived. Together with the increasing computational power of computers, machine learning is going to be a logical method to resolve this problem. Our main aim for this paper is to develop a much better stock price prediction system that the investments increase and therefore the stockholders can maximize their profits.

**PREDICTION METHODS:** 1. Fundamental analysis estimates securities by attempting to measure their native value. It is a method of determining a stock's real or "fair market" value. If the fair market value is higher than the market price, the stock is considered to be underestimated and a buy guidance is given. 2. Technical analysis seeks to predict future price movements, providing merchants with the facts needed to make a profit. Merchants appeal technical analysis tools to charts to recognize entry and exit points for possible trades.

## II. LITERATURE SURVEY

M. Sreemalli, P. Chaitanya, K. Srinivas [1] proposed the work of artificial neural network is a very approved technique to predict the stock market price and support vector machines. Using these models to list the pros and cons of all these models and compare the execution of stock market. Artificial Neural Network (ANN) appears great potentiality on machine learning problems such as classification and prediction. Using the linear model to execute the nonlinear class partitions, which has been occurred through some nonlinear mapping in which the input vector is fed into the high dimensional feature space. ARIMA model deals with time series data. In this paper, the use of machine learning techniques like Support Vector machine, Artificial Neural network and Auto Regressive Integrated moving average for the prediction of Nifty bank data. Dataset used here is 2015 Nifty bank dataset. Implementation using neural network takes more amount of time to perform computations compared to other techniques, whereas support vector machine has more error rate. Combining artificial neural network and Genetic algorithm might result in high accuracy. Indu Kumar, Kiran Dogra, Chetna Utreja, Premlata Yadav [2] proposed machine learning techniques have been applied for the stock price prediction in order to overcome such difficulties. In this paper, five models have been developed and their execution are compared in forecasting the stock market trends. The five supervised learning techniques are Support Vector Machine (SVM), Random Forest, K-Nearest Neighbor (KNN), Naïve Bayes, and SoftMax. The probing results show that Random Forest algorithm acts the best for larger datasets and Naïve Bayesian Classifier is the best for smaller datasets. The proposed architecture for the implemented work mainly consists of four steps: the feature extraction for the given dataset, supervised classification of the training dataset, supervised classification of the test dataset, and result evaluation.

Venkata Sasank Pagolu, Kamal Nayan Reddy Challa, Ganapati Panda [3] proposed the work to observe how well the changes in stock prices of a company, the rises and falls, are correlated with the public opinions being revealed in twitter about that company. Appreciate author's judgement from a piece of text is the purpose of sentiment analysis. In this paper, they have applied sentiment analysis and supervised machine learning principles to the tweets pulled out from twitter and examine the relationship between stock market movements of a company and sentiments in twitter. In a detailed way, positive news and tweets in social media about a

company would definitely uplift people to invest in the stocks of that company and as a result the stock price of that company would expand. At the end of the survey, it is shown that a strong association exists between the rise and falls in stock prices with the public sentiments in tweets.

Mariam Moukalled, Wassim El-Hajj, Mohamad Jaber [4] proposed an automated trading system that integrates mathematical functions, machine learning, and other external factors such as news' sentiments for the purpose of accomplishing better stock expectation precision and giving productive exchanges. Particularly, they aim to determine the price or the trend of a certain stock for the coming end-of-day considering the first several trading hours of the day. To achieve this goal, they trained traditional machine learning algorithms and created/trained multiple deep learning models taking into consideration the importance of the relevant news. Various experiments were conducted, the highest accuracy (82.91%) of which was achieved using SVM for Apple Inc. (AAPL) stock. They trained different models and compared the effectiveness of recurrent neural network (RNN), feed forward neural network (FFNN), support vector machines (SVM) and support vector regression (SVR) in predicting the direction of today close price with respect to yesterday close price based on the features. In this paper, they developed a stock price trend prediction system. To build these models they have gathered data from two sources (i) Historical stock market data from Reuters and (ii) news sentiment released on a certain stock; was gathered for 4 unique stocks more than 10 years.

Osman Hegazy, Omar S, Soliman and Mustafa Abdul Salam[5] proposed a machine learning model to predict stock market price. The proposed algorithm combines Particle swarm optimization (PSO) and least square support vector machine (LS-SVM). The PSO algorithm is used to optimize LS-SVM to predict the day-to-day stock prices. Proposed model is based on the work of stocks historical data and technical indicators. PSO algorithm choose best free parameters combination for LS-SVM to avoid over-fitting and local minima problems and upgrade forecasting accuracy. The proposed model was revealed and estimate using thirteen benchmark financials datasets and differentiate with artificial neural network with Levenberg-Marquardt (LM) algorithm. The proposed LS-SVM-PSO model convergence to the global minimum. Also, it is efficient to overcome the over-fitting problem which establish in ANN, mostly in case of fluctuations in stock sector. PSO-LS-SVM algorithm parameters can be tuned easily. The presentation of the proposed model is better than LS-SVM and differentiate algorithms. LS-SVM-PSO realize the lowest error value accompany by single LS-SVM, while ANN-BP algorithm is the trash one.

Ishita Parmar, Navanshu Agarwal, Sheirsh Saxena [6] proposed on the use of Regression and LSTM based Machine learning to predict stock values. Attributes considered are open, low, high, close and volume. The dataset should be as concrete as possible because a little change in the data can make massive changes in the outcome. In this paper, supervised machine learning is employed on a dataset obtained from Yahoo Finance. For the reason of simulation and examination, the data for only one company was examined. In this paper, both the techniques have shown an improvement in the accuracy of predictions, thereby yielding positive results with the LSTM model proving to be more efficient.

Aparna Nayak, M. M. Manohara Pai and Radhika M. Pai [7] proposed a try is made for prediction of stock market trend. Two models are construct one for daily prediction and the other one is for monthly prediction. Supervised machine learning algorithms are used to create the models. As part of the daily prediction model, historical prices are merged with sentiments. Up to 70% of accuracy is distinguish using supervised machine learning algorithms on daily prediction model. Monthly prediction model tries to estimate whether there is any resemblance between any two months trend. Sentiment from social media data and news are pull out. Extracted sentiments later will be integrated with historic price to construct prediction model. Two different models have been constructing to predict stock market trend.

T. Manojlović and I. Štajduhar[8] suggest 5-days-ahead and 10-days-ahead predictive models which are built using the random forests algorithm. The models are built on the historical data of the CROBEX index and on a hardly any companies listed at the Zagreb Stock Exchange from various sectors. Some technical indicators, popular in quantitative analysis of stock markets, are selected as model inputs. The suggested method is estimated using 10-fold cross validation, achieving an average classification accuracy of 76.5% for 5-days-forward models and 80.8% for 10-forward models. There are two common ways of predicting stock market behavior. The first one is formed on the prediction of future price values of a stock. This approach usually requires treating the historical data as time series data, feeding the distinct time frame signals to an algorithm and trying to model the future time points in the signal, e.g. Using ARIMA models. The second is based on forecasting the future price direction of a stock, i.e., calculating whether the price will rise or fall the next day, or in a couple of days.

Kunal Pahwa and Neha Agarwal [9] proposes to use machine learning algorithm to forecast the future stock price for interchange by using open-source libraries and pre-existing algorithms to assist make this uncertain format of business a little more certain. The outcome is entirely based on numbers and assumes a lot of principle that may or may not follow in the real world so as the time of prediction. They will be using the simplest classifier i.e., Linear Regression. the classifier perceives the features and simply looks at its label and recollect it. Finally, the result is plotted in graph. The key component of every consequence is the accuracy it conveys.

Radu Iacomin[10] proposed a machine learning algorithms such as ANN (artificial neural networks) with different feature selection. The results of this study will show that the algorithm of classification SVM (Support Vector Machines) with the help of feature selection (Principal component analysis) PCA will have the success of making a profit. The main goal is to detect the direction and price of 8 January 2014 and to validate the algorithms for a real prediction. In this paper they described results for the foreign exchange stock market using the two algorithms. For the GASVM algorithm the Rate of Recognition, ROR is 55%, meaning that only this amount of decisions is going to generate profit. For the PCASVM algorithm, which has a slightly increased ROR of 68%. PCASVM was applied to both

remove the false predictions and to govern what features are important. Differentiate the simple methods from SVM and developing to GASVM and PCASVM, the solution to the main problem and sub-issues is more systematic and shows promising results for a real forecasting using recent data sets.

Sumeet Sarode, Harsha G. Tolani, Prateek Kak, Lifna C S[11] proposed an approach combining two distinct fields for analysis of stock exchange. The system integrates price prediction found on historical and real-time data along with news analysis. LSTM (Long Short-Term Memory) is used for forecasting. The results of both analyses are integrated together to get a reply which gives a guidance for future increases. with the help of LSTM network, a very high level of accuracy can be achieved in predicting the upcoming trends and the price estimates of various stocks. To get the sentiment of every good in news/tweet its polarity has been found. Polarity of any news/tweet can be calculated by the two available approaches- the dictionary-based approach, and the semi-supervised algorithm. And this whole process gives a summarized result based on the sentiments of the News. On the other hand, the historical and real

time stock market prices are collected with the help of NSE (National Stock Exchange) Tools and further processed with the help of LSTM Model to predict the upcoming trends and prices. And the predicted results are then integrated with the summarized results. Thus, this paper proposes a system that would provide recommendations for the buying of shares of distinct companies.

Vaishnavi Gururaj, Shriya V R and Dr. Ashwini K [12] proposed Linear Regression (LR) which is a basic technique by which a linear trend can be obtained. In any case, Support Vector Machines (SVMs) have progressed highlights like high accuracy and consistency. In this paper, they survey the pros and cons of using both these to forecast values and differentiate both algorithms., with the help of RStudio development environment, this survey has been performed using a statistical language such as R. In this survey, they make use of exactly 1 year of stock data of The Coca-Cola Company, from January 2017 to 2018. Here a Simple Linear Regression in One Variable is examined, namely the Closing Stock price or the End of Day price for the prediction. The performance measures that were used to assess the predictive accuracy of the proposed system included, the mean absolute error (MAE), the mean square error (MSE), the correlation coefficient (R), the non-linear regression multiple correlation coefficient, the mean absolute percentage error (MAPE), root mean square error (RMSE).

K. Hiba Sadiya, Aditya Sharma, Adarrsh Paul, Sarmistha Padhi, Saurav Sanyal [13] proposed the best model to predict the value of the stock market. During the process of considering various techniques and variables that must be taken into account, they found out that techniques like random forest, support vector machine. In this paper they are going to present and review a more feasible method to predict the stock movement with higher accuracy. In this proposed system, they were able to train the machine from the various data points from the past to make a future prediction. They took data from the previous year stocks to train the model. The paper demonstrates the machine learning model to predict the stock value with more accuracy as compared to previously implemented machine learning models.

C.H. Vanipriya and K. Thammi Reddy [14] proposed model combines both of these methods to develop a hybrid Machine learning Stock Market Predictor based on Neural Networks, with an intent of improving the accuracy. This research is aimed at improving the efficiency of stock market prediction models by combining historical pricing models with sentimental analysis by developing a hybrid neural network to which historical prices and sentimental values are fed as inputs. There are mainly two methods for extraction of sentiments: Lexicon methods and machine learning methods. The system uses supervised training set with max 1000 iterations, learning rate 0.7 and Max\_error is 0.0001. Training stops after some iteration with a total net error under 0.0001.

Edgar P. Torres P, Myriam Hernández-Álvarez, Edgar A. Torres Hernández, and Sang Guun Yoo [15] proposed the probabilities of making prediction of stock market prices using authentic data and machine learning algorithms. They have experimented with stock market data of the Apple Inc. using random trees and Multilayer perceptron algorithms to execute the predictions of closing prices. They have used the historical data of the highest price, the lowest price, opening price, closing price, and volume traded of Apple Inc. stocks which has been assembled using Google Finance. Likewise, they have selected to predict the variable Close because it's the last price at which all members agreed for the financial security in the traded session. By extension, this method depends heavily on the sentiments, feelings, and expectations of participants. They have used the WEKA packages.

### III. COMPARATIVE STUDY OF VARIOUS ALGORITHMS IN LITERATURE REVIEW

TABLE I. COMPARATIVE STUDY OF DIFFERENT ALGORITHM

YEAR	AUTHOR	PURPOSE	METHOD MENTIONED	INFERENCE
2019	M. Sreemalli, P. Chaitanya & K. Srinivas	Comparative Analysis of Machine Learning Techniques on Stock Market Prediction	ANN, SVM and ARIMA	ARIMA is more efficient than ANN and SVM
2017	Indu Kumar, Kiran Dogra, Chetna Utreja & Premlata Yadav	A Comparative Study of Supervised Machine Learning Algorithms for Stock Market Trend Prediction.	SVM, Random Forest, KNN, Naive Bayes, and SoftMax	SVM has higher Accuracy on Small Dataset and large dataset
2016	Venkata Sank Pagolu, Kamal Nayan Reddy Challa & Ganapati Panda	Sentiment Analysis of Twitter Data for Predicting Stock Market Movements	Sentiment Analysis and Supervised Machine Learning Principles	At the end of the survey, it is shown that a strong association exists between the rise and falls in stock prices with the public sentiments in tweets
2017	Mariam Moukalled, Wassim El-Hajj, Mohamad Jaber	Automated Stock Price Prediction Using Machine Learning	SVM and other external factors such as news sentiments	By using SVM and news sentiments it gives more accuracy

2013	Osman Hegazy, Omar S, Soliman and Mustafa Abdul	A Machine Learning Model for StockMarket Prediction	Particle swarm optimization (PSO)and Least square SVM	The performance is better than LS-SVM and compared algorithms. LS-SVM- PSO accomplishes the least mistake esteem followed by single LS- SVM.
2018	Ishita Parmar, Navanshu Agarwal& Sheirsh Saxena	Stock Market Prediction Using Machine Learning	Regression and LSTM	It yields positive results with The LSTM model proving to be more efficient
2016	Aparna Nayak, M. M. Manohara Pai and Radhika M. Pai	Prediction Models for Indian Stock Market	Supervised machine learning algorithms and sentiments	70% of accuracy is distinguish using supervised machine learning algorithms on daily prediction model
2015	T. Manojlović and I. Štajduhar	Predicting Stock Market Trends Using Random Forests: A Sample of The Zagreb Stock Exchange	Random Forest	Experimental results suggest that random forests achieve the performance accuracy in stock market trend prediction.
2019	Kunal Pahwa and Neha Agarwal	Stock Market Analysis Using Supervised Machine Learning	Linear regression	Adv: Easier to implement, interpret and efficient to train
2015	Radu Iacomin	Stock Market Prediction	SVM, GASVM & PCASVM	PCASVM gives more accuracy as compared to SVM &GASVM
2019	Sumeet Sarode, Harsha G. Tolani, Prateek Kak, Lifna C S	Stock Price Prediction Using Machine Learning Techniques	LSTM	It is well-suited to process, classify and forecast time series given time lags of unknown time span
1934	Vaishnavi Gururaj, Shriya V R and Dr. Ashwini K	Stock Market Prediction Using Linear Regression and Support Vector	SVM &LR machines	As compared to LR, SVM have advanced features such ashigh accuracy and predictability
2019	K. Hiba Sadia, Aditya Sharma, Adarsh Paul, SarmisthaPadhi, &Saurav Sanyal	Stock Market Prediction Using Machine Learning Algorithms	Random Forest, Support Vector Machine	Best suitable algorithm for predicting the market price of a stock is the random forest algorithm
2016	C.H. Vanipriya1 and K.Thammi Reddy	Indian Stock Market Predictor System	Hybrid Neural Network	Here the error rate is less as compared to Machine learning



#### IV. CONCLUSION

Stock market prediction is the act of trying to regulate the future value of a company stock or other financial instrument traded on an exchange. The fortunate prediction of a stock's succeeding price could yield remarkable profit. From the literature survey, we found that the most suitable algorithm for forecasting the market price of a stock from the historical data based on various data points is the support vector machine. This algorithm will be a great benefit for dealers and stakeholders for investing money in the stock market since it is trained on a huge collection of historical data and has been selected after being drilled on a sample data

#### REFERENCES

- [1] Venkata Sasank Pagolu, Kamal Nayan Reddy Challa, Ganapati Panda “Sentiment Analysis of Twitter Data for Predicting Stock Market Movements”
- [2] Mariam Moukalled, Wassim El-Hajj, Mohamad Jaber “Automated Stock Price Prediction Using Machine Learning”
- [3] Osman Hegazy, Omar S, Soliman and Mustafa Abdul Salam”A Machine Learning Model for Stock Market Prediction”
- [4] Ishita Parmar, Navanshu Agarwal, Sheirsh Saxena” Stock Market Prediction Using Machine Learning 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC)
- [5] Aparna Nayak, M. M. Manohara Pai and Radhika M. Pai” Prediction Models for Indian Stock Market”
- [6] T. Manojlović and I. Štajduhar “PREDICTING “Stock Market Trends Using Random Forests: A Sample of The Zagreb Stock Exchange” MIPRO 2015, 25-29 May 2015, Opatija, Croatia
- [7] Kuna Pahwa and Neha Agarwal “Stock Market Analysis Using Supervised Machine Learning”
- [8] Radu Iacomin “Stock Market Prediction”2015 19th International Conference on System Theory, Control and Computing (ICSTCC), October 14-16, Cheile Gradistei, Romania
- [9] Sumeet Sarode, Harsha G. Tolani, Prateek Kak, Lifna C S “Stock Price Prediction Using Machine Learning Techniques” International Conference on Intelligent Sustainable Systems (ICISS 2019)
- [10] Vaishnavi Gururaj, Shriya V R and Dr. Ashwini K “Stock Market Prediction Using Linear Regression and Support Vector Machines” <http://www.ripublication.com/>
- [11] K. Hiba Sadia, Aditya Sharma, Adarrsh Paul, SarmisthaPadhi, Saurav Sanyal “Stock Market Prediction Using Machine Learning Algorithms” International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-8 Issue-4, April 2019
- [12] C.H. Vanipriya I and K. Thammi Reddy “Indian Stock Market Predictor System”
- [13] Edgar P. Torres P, Myriam Hernández-Álvarez, Edgar A. Torres Hernández, and Sang Guun Yoo “Stock Market Data Prediction Using Machine Learning Techniques”

