

# ANALYSIS ON STOCK MARKET PREDICTION USING MACHINE LEARNING TECHNIQUES

*Sheikh Irfan Akbar*

*Research Scholar, School of Engineering and Technology, RIMT University, Mandi Gobindgarh, Punjab, India 147301*

## Abstract

***Stock value prediction is one in every of the foremost wide studied and difficult issues that attracts researchers from several fields together with political economy, history, finance, arithmetic, and computing. The volatile nature of the exchange makes it troublesome to use easy time-series or regression techniques. The efficient prediction algorithm gives the traders more benefit. The accuracy of prediction models depends on the features selected which will be the input to the prediction model. The results suggests that the data captured from non-traditional experts like google and Wikipedia and the data captured from desperate data sources can improve the performance of financial prediction model. In this paper I have analyzed some stock prediction models. I have also done the survey on stock prediction models.***

***Keyword: stock market, nifty, Sensex, Artificial neural networks, stock market efficiency and neural network, Efficient Market Hypothesis, SVM, and National Stock Exchange.***

## 1. INTRODUCTION

Stock market price movement is one of the most emerging fields in the field of computer science, economics, finance, and management, etc. Stock market one day ahead has been a goal of many traders. The question arises here: "To what extent can the history of a common stock's price be used to make meaningful predictions concerning the future price of the stock. However, information concerning a stock is normally uncertain, vague and incomplete making it a dare to predict the future economic performance. Accurate

stock market predictions are very important. The most among all is the need for investors, to prevent against potential market risks and the opportunities to make profits by trading indexes. Traditional analysis methods reliability is strongly relying on experience, is somewhat being doubted due to the complication of correlated information, and the mode of stock exchange. The successful prediction, of a stock's future price, could yield a considerable profit. The stock market is essentially a nonparametric and non-linear system that is extremely hard to model with any absolute accuracy. The researchers said that fundamental analysis assumes that the price of a stock depended on its intrinsic value and expected a return on investment. Analyzing the company's operations and the market in which the company is operating, the stock price can be predicted reasonably well [Qasem A. Al-Radaideh, 2013](#). Most people believe that fundamental analysis is a good method only on a long-term basis. However, for short- and medium term speculations, fundamental analysis is not suitable.

Indian exchange has gained the interest of investor's investment in 2 main exchange named as Mumbai exchange (BSE) and National exchange (NSE). There is high risk concerned for investors as a result of a lot of quality of the exchange. The Sensex and bang-up area unit 2 such outstanding market indices that perform at intervals the Indian exchange. These 2 market indexes represent the stocks for animal disease (Bombay Stock Exchange) and NSE (National Stock Exchange) severally. Specifically, below animal disease there area unit thirty corporations for Sensex, whereas below NSE there area unit fifty corporations for bang-up therefore it's known as NIFTY50. Therefore there ought to predict the

exchange standing for investors by exploitation these 2 most significant indicators i.e. Sensex and bang-up.

## 2. STOCK MARKET EFFICIENCY

In keeping with the efficient market hypothesis (EMH) Bin Weng a, 2017 stock costs can solely reply to new data then can follow a stochastic process. If they solely reply to new data, they cannot be expected. That the stocks follow a stochastic process is a symbol of market potency since predictable movement would mean the market costs weren't reflective that data. There square measure three variants of this theory – weak, semi-strong, and strong. Most analysis has ended that the semi-strong version holds true. This version claims that stock costs replicate all in public offered data, however non-public data may be accustomed below the belt predict profits. This is often the premise behind sturdy trading laws. Notwithstanding, bound market phenomena truly run contrary to EMH. These square measure called market anomalies. Contradicting the efficient market hypothesis, stocks that have recently been increasing still increase, and recently decreasing stocks still decrease. This kind of trend implies some quantity of certainty to future stock costs. If the EMH command utterly true, then the direction of future stock costs couldn't be expected with bigger than five hundredth accuracies. That is, one mustn't be ready to guess whether or not future costs can go up or down higher than the easy random estimate.

## 3. MACHINE LEARNING METHODS

### 3.1. artificial neural network (ann)

The Artificial Neural Network (ANN) has many benefits however one amongst the foremost recognized of those is that the proven fact that it will truly learn from perceptive information sets. During this method, Artificial Neural Network is thought for as a random operate approximation tool Milind Kolambe, 2016. These styles of tools facilitate estimate the foremost efficient and ideal ways for inbound at solutions whereas shaping computing functions or distributions. ANN uses information samples rather than complete information sets to reach solutions, which saves each time and cash. ANNs area unit thought-about fairly easy mathematical

models to boost the effectiveness of the out there the information analysis technologies. ANNs includes three layers. These layers area unit connected to every different. The primary layer consists of input neurons. Those neurons send information on to the second layer that successively sends the neurons to the third layer. Coaching a man-made neural network involves selecting from allowed models that there are many associated algorithms. Phua et al. 2003 applied Neural Networks to the money prediction. He tested the influence of volume information on Stock value prediction. Khan et al. applied the Neural Networks with completely different range of hidden layers (as shown in fig1) to research the prediction of the Stock costs.

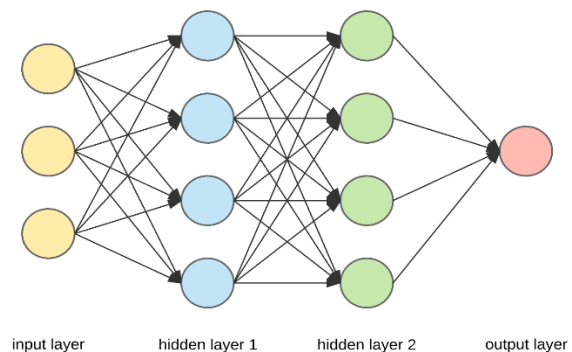


Fig1 ANN

### 3.2. support vector machines

Although Support vector machines (SVM) may be applied to numerous improvement issues like regression, the everyday downside is to classify the information. The essential plan is shown in the figure. The information point's square measure known as being positive or negative and the problem are to search out a hyperplane. This hyperplane separates the points (data) by a top margin.

“Support Vector Machine” (SVM) falls below a class of supervised machine learning algorithms which might be used for each classification and regression challenges. However researchers mostly use it for classification issues. In this algorithmic program, we tend to plot every information item as a degree in n-dimensional area (where n is range of options you have) with the value of every feature being the worth of a selected coordinate

[4]. Then, we tend to perform classification by finding the hyperplane that differentiates the two categories SVM (support vector machines) and ARIMA (Autoregressive Integrated Moving Average) model and located it gave promising results.

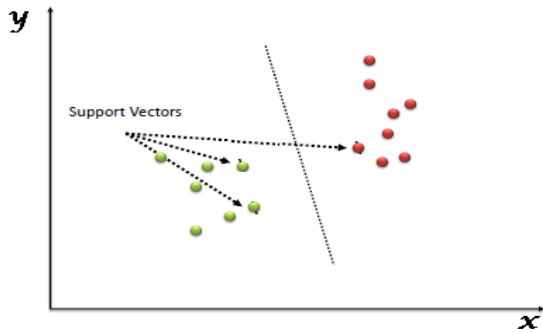


Fig 2: support vector machine

### 3.3. decision trees

Decision tree can be defined as a decision support tool which builds classification or regression models within the kind of a tree structure. It breaks down a dataset into small then even tinier subsets whereas at an equivalent time Associate in nursing associated call tree is incrementally developed. The final result could be a tree with call nodes and leaf nodes [4]. A call node has two or a lot of branches. Leaf node represents a classification or call. The uppermost call node during a tree that corresponds to the simplest predictor called root node. Call trees will handle each categorical and numerical information. A call tree is made top-down from a root node and involves partitioning the info into subsets that contain instances with similar values. The information gain is predicated on the decrease in entropy once a dataset is split on Associate in the nursing attribute. Constructing a call tree is all concerning finding attribute that returns the best info gain (i.e., the foremost homogenous branches). Nair B. B. et al. 2010 projected a system supported a genetic algorithmic rule optimized call tree-support vector machine hybrid, which may predict one-day-ahead trends available markets

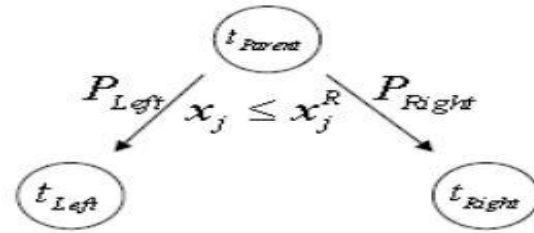


Fig3 Decision trees

## 4. LITERATURE SURVEY

Bhardwaj, A. et, al. 2015 conducted a study on Sentiment Analysis for Indian Stock Market Prediction Using Sensex and Nifty. The main focus of the research work was importance of sentiment analysis for stock market indicators such as Sensex and Nifty to predict the price of stock. For implementation purpose the proposed system fetched the live Sensex and Nifty data values from Timesofindia.com. Python script was run with sleep count time interval of one second for fetching the data, and values were calculated for different time interval. After that result is drawn which shows that for a particular time interval the fetched values of Sensex and Nifty remains constant. It was proposed that we should use python scripting language which have fast execution environment and this will help out the investors in order to make a prediction of on what shares money should be invested, it will also help in maintaining the economical balance of share market

Stock market prediction is defined as the act of trying to provide the future price movement of a company's stock or other financial instruments traded on a financial exchange. Shweta Tiwari and Rekha P 2013 predicting future trends in the stock market by the decision tree rough-set based hybrid system with HHMM. In this paper, researchers have developed a hybrid system based on the decision trees for predicting the future trends in the Bombay Stock Exchange (BSE SENSEX) with the combination of Hierarchical Hidden Markov.

Data mining technology was used to stock market to research the trend of price. Bhagwant Chauhan, Umesh B and Sachin Kale 2014 have proposed a model on Stock Market Prediction Using Artificial Neural Networks. The study aims to predict the future trend of the stock

market and the fluctuation of price. The findings of the study say that ANN has shown to be effective approach for pattern recognition, classification, clustering and especially time series prediction with a great degree of accuracy. The aim of this work was the implementation of neural networks with back propagation algorithm for the stock market. Nevertheless, their performance is not always satisfactory. Back propagation algorithm is the best algorithm to be used in Feed forward neural network because it reduces an error between the actual output and desired output in a gradient descent manner.

Deep learning neural networks has the ability to extract features from a large set of raw data. This ability is hidden from predictors, which makes it potentially attractive for stock market prediction Eunsukh Chong, chulwoo Han, 2017. They provide a comprehensive and objective assessment of both the advantages and drawbacks of deep learning algorithms for stock market analysis and prediction in the paper. The study was carried out with the intraday stock data returns as input data and the extraction of features by three unsupervised methods principal component analysis, auto encoder, and the restricted Boltzmann Machine on the network's overall ability for the prediction of future market behaviour. The researchers evaluated prediction performance using four measures: normalized mean squared error (NMSE), root mean squared error (RMSE), mean absolute error (MAE), and mutual information (MI). Based on above grounds the researchers were able to find that the DNNs perform better than a linear autoregressive model in the training set, but the advantage mostly disappears in the test set.

Kartik Sharma, Akhelesh Rao, 2017 conducted a study on Stock Market Analysis. The proposed system was an attempt to reconcile computed sentiments alongside traditional/more common data mining. This will be accomplished with the help of 2 types of datasets. Firstly, historical data from Google Finance will be mined to garner traditionally available predictions. Two independent predictions, are then combined to generate a final output which will be used to predict the next day's opening price. Datasets consisting of

historical data as well as recent headlines were mined to ascertain stock price movement. This main aims of the system was to predict stock price movement more accurately by emulating instinctual reasoning by implementing sentiment analysis. It helped the proposed system ascertain sentiment analysis as the better companion to the traditional data mining approach instead of employing a neural network in cases which called for the supervised approach. It was noted however that a neural network worked extremely well in situations which called for the unsupervised approach. The research concludes that the unsupervised and supervised learning depend methods help to find the results in a better way. The combinational study was done to get better accuracy. Further, optimizations can be done in sequence to get improved results.

The model on Stock market one-day ahead movement prediction using disparate data sources was proposed to evaluate the performance of expert system, the researchers present a case study based on the AAPL (Apple NASDAQ) stock. They believed that their expert system had an 85% accuracy in predicting the next-day AAPL stock movement, which outperforms the reported rates in the literature. To predict stock movements, the researchers Bin Weng, Mohamed A. Ahmed and Fadel M. Megahed, 2017 propose a data-driven approach that consists of three main phases. They also populated additional features (i.e. summary statistics) in an attempt to uncover more significant predictors for stock movement. Based on the evaluation, they select an appropriate model for real-time stock market prediction. The results of the study suggest:

Sentiment analysis for stock market indicators such as Sensex and Nifty to predict the price of stock has a very big impact. The research on Sentiment Analysis for Indian Stock Market Prediction Using Sensex and Nifty by Aditya Bhardwaj, Yogendra N 2015. They proposed a model to fetch live server data by using Python programming language, which is used for performing sentiment analysis on the extracted datasets from online news portal. For implementation purpose the proposed system fetched the live Sensex and Nifty data values from Timesofindia.com. Python script was run with sleep

count time interval of one second for fetching the data, and values were calculated for different time interval. After that result is drawn which shows that for a particular time interval the fetched values of Sensex and Nifty remains constant. It was proposed that we should use python scripting language which have fast execution environment and this will help out the investors in order to make a prediction of on what shares money should be invested, it will also help in maintaining the economical balance of share market.

## 5. CONCLUSION

In this paper I have analyzed various machine learning algorithms for stock market prediction. With the increase in technology and the use of artificial intelligence the efficiency of prediction models have increased. In this paper I have studied about the three basic machine learning models, artificial neural networks, Decision trees and Support vector machines. From the analysis it is clear that ANN and SVM are mostly used. Adding some features to the historical stock prices increases the efficiency of the prediction model.

## 6. ACKNOWLEDGEMENT

I would like to thank Almighty Allah for giving me this opportunity and my parents for their support and courage at every step.

I also extend my word of thanks to all the authors and researches whose work and research was really helpful.

## REFERENCES

- [1] Bhagwant Chauhan, Umesh Bidave, Sachin Kale "Stock Market Prediction Using Artificial Neural networks," *International Journal of Computer Science and Information Technologies*, vol. 5, 2014.
- [2] QASIM A.AI-RADAIDEH, ADEL ABU ASSAF "predicting stock prices using data mining techniques," *International Arab Conference on information Technology*, 2013.
- [3] S. Madge, "Predicting Stock Price Direction using Support Vector machines," 2015.
- [4] M. Kolambe, "Survey Paper on Stock Market Prediction," *International Journal of Innovative Research in Computer and communication engineering*, vol. 4, no. 10, 2016.
- [5] Neha Khalfay. Deepali Vohra, Vidhi Soni, Nirbhey Singh Pahwa, "Stock Prediction using Machine Learning a Review Paper," *International Journal of Computer Applications*, vol. 163, no. 5, 2017.
- [6] Mohamad A.Ahmad, Fadel M Megahed. Bin Weng, "Stock market one-day ahead movement prediction using disparate data sources," *Expert Systems With Applications*, pp. 153-163, 2017.
- [7] Kartik Sharma,Akilesh Rao, Sujay sail "STOCK MARKET ANALYSIS," in *6th international conference on latest inovations in science and engineering*, new delhi, 2017.
- [8] Aditya Bhardwaj, yogendre narayan, vanraj "Sentiment Analysis for Indian Stock Market Prediction Using Sensex and Nifty," in *4thInternational Conference on Eco-friendly Computing and Communication Systems*, Chandigarh, 2015.