© UIJIR | ISSN (O) – 2582-6417
SEP. 2020 | Vol. 1 Issue 4
www.uijir.com

Universe International Journal of Interdisciplinary Research
(Peer Reviewed Refereed Journal)

A SURVEY ON STOCK MARKET PREDICTION USING MACHINE LEARNING TECHNIQUES

Author Name: Muhammad Zeeshan Younas

Affiliation: *Department Of Computer Science, Capital University of Science & Technology, Islamabad, Pakistan.*

E-Mail: Itszeeshanyounas@gmail.com

Abstract

Stock market prediction has been a field of attention for researchers as well as investors. The major task Stock Market prediction is a big challenge and difficulties stage for finance and statistics experts. This paper suggests a methodology to prediction of stock market current developments using machine learning models and techniques like Random Forest model and Support Vector Machine. Stock exchange prediction is not easy but machine learning with some techniques and method have solved some problem for future prediction. Different steps are followed to apply and get best stock market predictions. This research work with advanced techniques that is done by various researchers was studied in this research work it provides admirable and good accuracy performance by perceiving different research papers. SVM is a satisfactory tool for economic forecasting and stock market prediction.

Keywords *: Machine Learning, Classification Techniques, Data Mining, Stock Market, Prediction*

INTRODUCTION

Stock market is a centralized place where the companies publicly can trade with their shares, they can buy and sell their shares at any time. Stock exchange allows companies to increase their capital of the business and makes stock shares available to publicly for buying. Stock market offer two various important purposes, the first one is provide capital to different companies for expanding their businesses and on the other hand secondary purpose investor can purchase stocks and sell respective stocks whenever the price increases from the actual purchased price and some stocks pay regularly dividends. It is very easy way to earn money and increase capital of the business, but it is also a risky because investors never know about stock market crash, a well-known example is Covid-19. Everybody wants to earn more and more but major issue is risk in every business and stock market. “The major task Stock Market prediction is a big challenge and difficulties stage for finance and statistics experts, the core motive behind prediction is purchase the stocks that are probably to rise in stock market and sell stocks that are likely to fall. There are the following two methods for stock market prediction, first method is technical analysis that is based on earlier values and prices. And second one is fundamental

analysis that is based on company's techniques and basic information like market capitalization and annual growth rates" (Mojtaba,et al 2019). Mostly investors consider that fundamental analysis is a very good way only for a long term basis.

Stock market prediction is not easy because it may anytime goes to crash due to some natural disaster or others. The effective prediction of a stocks market's future price will increase investor's gains and interest. Investors are trying to find a good way to predict stock market prices and to find the factual stocks for time to sell or buy shares. If we use integrated techniques with the stock market past information system then it would be very beneficial and it will also minimize the risk in future. It can be done by comparing various machine learning techniques for discovering the appropriateness, ML algorithms for mining the hidden patterns from huge database and machine learning classification techniques combine statistical analysis.

Many researchers apply different machine learning and data mining classification models to help stock market prediction with improved accuracy level. Namely, Multiple Additive Regression Tree (MART), Multi-Layer Perceptron (MLP), Naïve Bayes, K-NN and SVM some techniques used here. Stock market was usually predicted by financial professionals in the earlier duration. Nevertheless data scientists have started resolving prediction difficulties with the improvement of machine learning techniques. Many researchers and scientists have initiated using ML techniques and method to increase the performance of prediction and improve the accuracy of predictions. "Prediction of stock market is not easy it is mostly a full of challenges face, many data scientist and researchers normally faced some problems and issues when they tried to develop a predictive model for stock market, Non-linearity and complexity are the two major challenges triggered by the uncertainty of stock market and the association among investment mindset and market performance" (Duarte, et al 2017). Prediction can be generate by different attributes to find whether this stocks will have worth in future or not, furthermore it takes minimum time for prediction and increase the stock exchange market value with good accuracy and decreases the occurrence of loss. Different steps are followed to apply and get best stock market predictions. Support vector Machine (SVM) is a satisfactory tool for financial forecasting. This research work with advanced techniques that is done by various researchers was studied in this research work. SVM is well-organized method for predicting Stock market.

TECHNIQUES AND METHODS

Mainly there are five techniques for the prediction of Stock Market:

i. Support-Vector Machine (SVM)
ii. Multi-Layer Perceptron (MLP)
iii. Multiple Additive Regression Tree (MART)
iv. Naïve Bayes (NB)
v. K-Nearest Neighbors (K-NN)

i. Support-Vector Machine (SVM):

SVM is a very common used supervised ML algorithm, it is used for linear and non-linear data. It is very effective in high dimensional space. It includes identifying hyperplanes which separate data into various classes. If hyperplanes are exposed then fresh data points can be simply classified. SVM is actually multipurpose and also clever of performing linear or nonlinear regression and classification. The objective is to find highest margin hyperplane that is uttermost from the nearest points in the various two classes, and this point is known as support vector. SVM works really good when it hold clear margin of separation and where amount of dimensions is higher than the amount of samples. But when data set is noisy its performance is very low and does not perform well, when it have large amount of data.

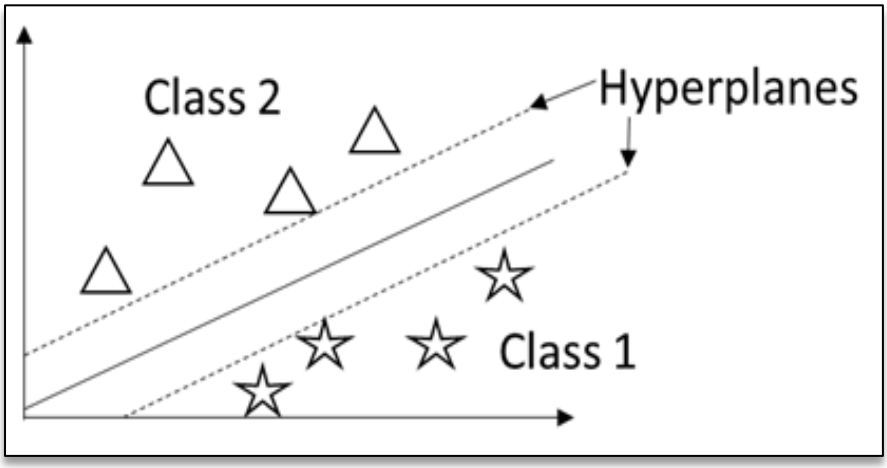

Figure 1 -Support-Vector Machine (SVM)

ii. Multi-Layer Perceptron (MLP):

It is a feed-forward artificial neural network (ANN) and it has one extra layer that is called hidden layer. Hidden layer may be one and more than one and this layer

holds the intermediary neurons. MLP uses the methods of back propagation for training its many layers and its major benefit is individual data that is not linearly distinguishable [3]. MLP is mostly used for handling problems that need supervised learning more likely parallel distributed processing and computational neuroscience. Application includes image recognition, speech recognition and machine translation.

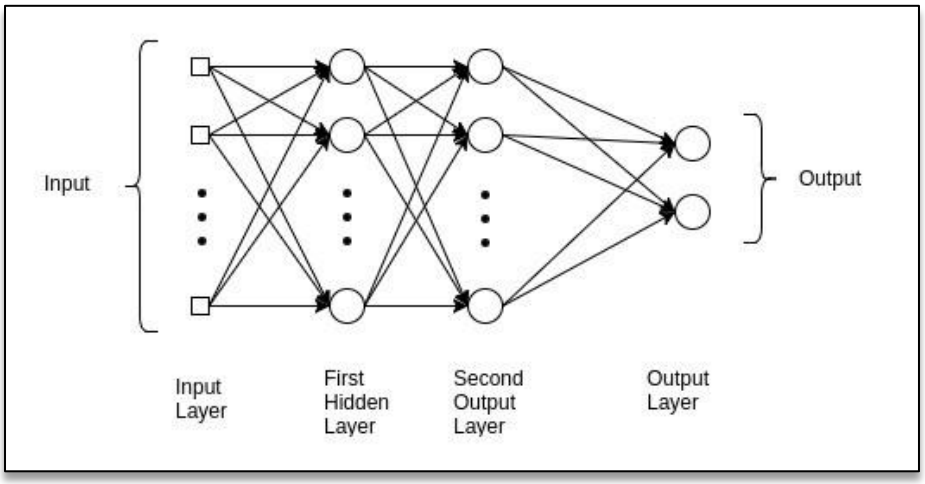

Figure 2 - Multi-Layer Perceptron (MLP)

iii. Multiple Additive Regression Tree (MART):

It is a new methodology mostly used to resolve prediction difficulties and problems that is based on huge datasets normally originate in Data mining and it is an method of gradient tree to boosting for predictive data mining (classification and regression). MART holds the very common and simple classification trees are fit. MART incline to be impervious to makeovers of predictors and response features, outliers, missing values, and to the presence of possibly huge numbers of inappropriate predictor variables that have little or no influence on the results.

iv. Naïve bayes:

A naïve bayes classifier is a supervised ML algorithm that is used for classify the specific data into predefined classes. Naïve bayes classifier make use Bayes' theorem and it has independence between data points and attributes. Most popular uses of naïve bayes text analysis, spam filter and medical diagnosis. It uses conditional probability to classify the test dataset. Naïve bayes requires

insignificant amount of training dataset and fast to predict the classes of test data. It does not appropriate for large data and its working is not good if the features are associated

v. **K-Nearest Neighbors (K-NN):**

It is a simplest ML algorithm that based on supervised learning methods. KNN is a famous method to classifying an invisible instance and it categorizes an innovative data point that is based on the resemblance. When we apply this classification on new data set it can be effortlessly classified that data into a fit suited range by using K-NN. This can be performed by using Euclidean distance. If we choose a right value then it will help for improving the accuracy of the respective model. K-NN implementation is very simple and it creates no erstwhile assumption for the data.

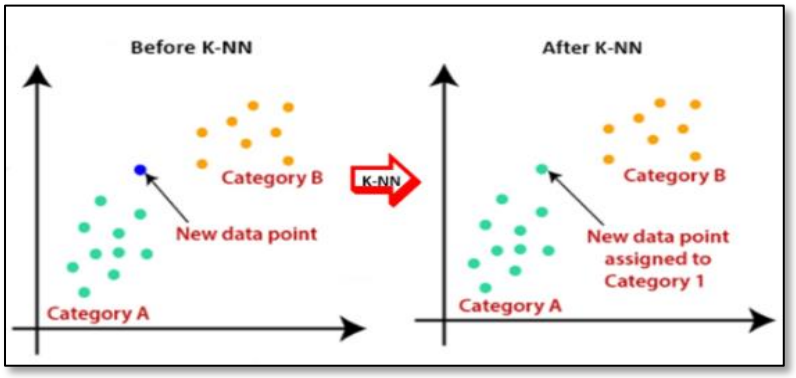

Figure 3–KNearest Neighbor (KNN)

LITERATURE REVIEW

Mojtaba Nabipour, et al (2020) describes about decrease the risk of development of prediction with deep learning and machine learning's algorithms, they proposed method is prediction of stock market changes. They have done prediction on four market group by using nine ML techniques and two deep learning techniques LSTM and RNN. Dataset is collected from historical records that contain ten technical features. They have claimed LSTM and RNN are the best approaches in following models and produced various result on different ML algorithms. This work is done to understand different pros and cons and provide results for various algorithms and their accuracy.

Wasiat Khan, et al (2020) they have compared different algorithms results to find a reliable classifier, and they have used deep learning classifier to achieving maximum prediction

results. They have claimed by their experimental results highest accuracy is 75.16%. Stock market 100% future prediction is not possible only correct decision making is best for purchased stocks. Stock traders should by respective stocks whose price is expected to increase in coming future and sell those stocks those are losing their values in stock exchange.

Shunrong Shen, et al (2012) proposed a method for prediction of stock market index movements using machine learning algorithms and data gathered from different global financial markets. They have conducted studies to predict the current stock market moments. In this study they have proposed a new prediction algorithm that working is related to chronological correlation between different financial product and global market to predict the upcoming stock trends in market. They have claimed some accuracy results by using various algorithms DJIA with 77.6%, S&PS500 76% and NADSDAQ with 74.4% accuracy. Furthermore they have used these same algorithms with different regression techniques to find actual addition in the stock markets.

Mehak Usmani, et al (2016) have done a work to predict stock market prediction and the core aim of this research is to confirm that ML techniques are clever of prediction the stock market performance routine in this era. Stock market prediction is not easy because it depends on current nature of stock market shares. They have done work by using different machine ML techniques and get various accuracy performance results. Best accuracy is claimed by this research work and that is 77% by using Multi-Layer Perceptron (MLP) algorithm. The results compared and the accuracy got as the followings: SLP 60%, MLP 77%, and Radial Basis Function (RBF) 62% and SVM 60% respectively.

Sahaj Singh Maini, et al (2017) this research work core aims to predict the way of stocks market trends in near future. They have discussed following two ML trends Support Vector Machine (SVM) and Random Forest Model (RFM) that is used to predict best accuracy results and a consistent prediction of stock market by using best historical global financial dataset. It is very difficult to analysis that can affect the decision making for buy and sell stocks in an optimistic manner. They have claimed accuracy compared results are, Random Forest Model (RFM) 84% and on the other hand Support Vector Machine (SVM) generate predictions with an accuracy result of 85.1%.

Kunal Pahwa, et al (2019) Stock market prediction is very sophisticated and complex way for business. Stock market first name is profit and second is risk. Everything is depends on future prediction. They have used various ML algorithms techniques to predict stock market future price, According to this research machine earning is well futures prediction tool we have seen

and till now, is a very dominant tool for prediction. There are different tool and techniques available to control and resolve various problems. This research work is restricted to only supervised ML

COMPARITIVE STUDY

Table 1: Accuracy comparison on Stock Exchange Dataset by various authors with proposed model

Technique or Methodology	Accuracy	Future scope	Advantages	Disadvantages
Naïve bayes	83.75	By their Experimental results they improved the models	It requires a minor amount of training data and predict to fast results	If the features are correlated its performance slow and not suitable for large data set
K-NN	61%	Its major focus on Find best stock exchange accuracy results	Robust for noisy data and simple implementation	Prediction time is very high as the others.
MART	73%	Focus on improving the prediction of different stock exchanges	Random forest can be used in MART it clarifies controlled and uncontrolled events	Huge numbers of inappropriate predictor variables that have little or no influence on the results

MLP	77%	Proposed a machine learning that can provide better accuracy in area of variety stock exchanges	Data driven and self-adaptive itself	Low speed of learning and large complexity of the network structures
SVM	85%	In future cooperative methods and techniques are applied to get more accuracy results	Produced accurate and better results even when input is linearly and non-linearly	Low performance if the data set is noisy

DISCUSSION AND CONCLUSION

The various Stock market prediction techniques and method have been discussed and analyzed is this research work. Machine learning and some data miming techniques are discussed. Many researchers apply different machine learning and data mining classification models to help stock market prediction with improved accuracy level. Mostly investors consider that fundamental analysis is a very good way only for a long term basis.

Stock market is a centralized place where the companies publicly can trade with their shares. Stock exchange prediction is not easy but machine learning with some techniques and method have solved some problem for future prediction. Different steps are followed to apply and get best stock market predictions. Support vector Machine (SVM) is a satisfactory tool for financial forecasting. This research work with advanced techniques that is done by various researchers was studied in this research work. At the end from the reasonable study we can easily conclude that Support Vector Machine (SVM) is well-organized method for predicting Stock market. It provides admirable and good accuracy performance by perceiving different research papers.

REFERENCES

1. 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 on the Tehran stock exchange.IEEE Access. 150199-150212, vol 8 , DOI 10.1109/access.2020.3015966.
2. Duarte, Juan Benjamin Duarte, Leonardo Hernán Talero Sarmiento, and Katherine Julieth Sierra Juárez (2017). Evaluation of the effect of investor psychology on an artificial stock market through its degree of efficiency. 62.4, 1361-1376.
3. Ching-seh (Mike) Wu, Mustafa Badshah and Vishwa Bhagwat (2019). Heart Disease Prediction Using Data Mining Techniques. 2nd International Conference on Data Science and Information Technology: 7–11, DOI 10.1145/3352411.3352413.
4. Wasiat Khan, Mustansar Ali Ghazanfar, Muhammad Awais Azam, Amin Karami, Khaled H.Alyoubi and Ahmed S. Alfakeeh (2020). Stock market prediction using machine learning classifiers and social media news. Journal of Ambient Intelligence and Humanized Computing, springer. DOI 10.1007/s12652-020-01839-w.
5. Shunrong Shen, Haomiao Jiang and Tongda Zhang (2012). Stock Market Forecasting Using Machine Learning Algorithms.http://cs229.stanford.edu.
6. Mehak Usmani, Syed Hasan Adil, Kamran Raza and Syed Saad Azhar Ali (2016). Stock Market Prediction Using Machine Learning Techniques. 3rd International Conference On Computer And Information Sciences (ICCOINS). DOI 10.1109/ICCOINS.2016.7783235.
7. Sahaj Singh Maini and Govinda.K (2017). Stock Market Prediction using Data Mining Techniques. Proceedings of the International Conference on Intelligent Sustainable Systems ICISS. DOI 10.1109/ISS1.2017.8389253.
8. Kunal Pahwa and Neha Agarwal (2019). Stock Market Analysis using Supervised Machine Learning. International Conference on Machine Learning, Big Data, Cloud and Parallel Computing COMITCon, DOI 10.1109/COMITCon.2019.8862225.