

Prediction of Stock Market Using Machine Learning Algorithms: A Review

Madhanraj¹, Maharishi², Madhan³, Jeevajothibalan⁴

^{1,2,3,4}Student, Department of Computer Science & Engineering, Sona College of Technology

Abstract

The study researches the usage of equipment learning algorithms in stock market prediction, emphasizing the significance of accurate prophesying for shareholders and financial viewers. It debates the occurrence of brainy networks in stock market prediction owing to their capability to grasp compound data inclinations and bonds, leading to more exact prophesy weighed against traditional methods. The research showcases the rise in publications related to machine learning applications in stock market prediction, with a peak in 2013 and continued interest in the subsequent years. Various machine learning tactics, including brainy networks, uphold vector machines, and Long-Term Memory (LSTM) models, are analyzed for their efficiency in delivering precise stock market guesses. The investigation aims to enlarge researchers' comprehension of the most suitable approaches for prophesying stock market motions by studying existing literature and models.

Keywords: Machine learning, LSTM, RNN, Stock Market, Prediction, AI

1. Introduction

To predict the stock market is quite a difficult thing to do. generally to predict the stock market there are two techniques.

- The technical method
- The fundamental method.

In the technical method, there are some patterns that are formed by using the past data in a particular stock. In the fundamental process, it is predicted using the company's background and depends on the investors. Since now these are the common old techniques used by people. Now the blooming techniques to predict the stock market are algorithms, Artificial intelligence and machine learning. Some of the highly used algorithms are listed below with accuracy.

1. SVM (Support Vector Machine)
2. RF (Random Forest)
3. NB (Naive Bayes)
4. RCAP (Regularized Class Association Rules Algorithm)
5. LSTM (Long Short-term Memory)
6. ANN (Artificial Neural Network.)

This Artificial intelligence uses all of the above mentioned algorithms. So, in this review paper, we can learn about each and every algorithm that are used with the help of Artificial intelligence and machine learning and can able to know about the accuracy of those algorithms.

2. GENERAL TYPES OF MACHINE LEARNING ALGORITHMS USED IN STOCK MARKET PREDICTION

Machine-learning way com is handy in forecasting market stock. (Nayak et all) have contrasted with some supervised machine-learning methods. Badlings et all have run with random forest (RF), AdaBoost, kernel factory, NN, SVM, and k-nearest neighbors (KNN) for forecasting the market stock trend for a whole year. (Patel et all).

<p>SUPPORT VECTOR MACHINE(SVM)</p>	<p>Support Vectors Machinery (SVM) will use two types of algorithms to predict the data. 1)Linear (Support Vectors Machinery) 2)Non-Linear (Support Vectors Machinery) The Linear (SVM) will separate the data into linearly.The Non-Linear (SVM) will separate the data into non-linearly. Using these two algorithms the (SVM) can run faster compared then the ANN(Artificial Neural Network). This Support Vector Machine has an unbelievable power to remember data. The SVM is commonly solve the problems in the process of ‘pattern and prediction and analyzes and map both linear and non-linear functions’.</p>
<p>RANDOM FOREST(RF)</p>	<p>Random Forrest (RF) is a group of classification or regression trees, that begin' made with guided sampling' of training data and randomly selected characteristics in tree induction.. Pre- dictions are made by aggregating (majority or average) the prophecies of the groups. Accord- ing to the forecasts of the decision tree, the ran- dom forest model decides the result. A higher number of trees generally means higher accura- cy. The best way to hurry up the random forest be to lessen the numbers of estimators. The best way to better the accurateness of the model be to up the number of trees. After training' the model, each tree in the forest be given a classi-</p>
	<p>said to vote for that class. RF can make regres- sion and classification tasks. RFs supply good forecasts that are easy to follow. It is able to manage large amounts of data efficiently. In comparison' to the decision tree algorithm, the RF algorithm supplies higher accurateness in prophesying' the result. Several researches have employed RF for stocks market prediction, in-</p>

	<p>cluding' Indu Kumar et al. [1], Mehar (Vijh et al). [3], M. Azmi, G. C Runger & A. Berrado [6], an' Mahla Nikou et al. [7].</p>
<p>NAÏVE BAYES(NB)</p>	<p>The fictive Baye (NF) being classifying algorithms using the Bay's theorem. It predicting the probability of belonging to every classes. The most probable class be the class highest probabilities. NF operating on categoric inputs variables rather of numeric variables. It be helpful for forecasting and predicting data founded on historical results. For immediate forecast, we consider NF as a quick- learning classifiers. Hence, it can using for creating immediate forecasts. Indira Kumar and oters [1] also performing testing on NB for stock market predictions!</p>
<p>INTERPRETABLE REGULARIZED CLASS ASSOCIATION RULES ALGORITHM FOR CLASSIFICATION(RCAP)</p>	<p>Creating high accuracy and understandable classifiers in space rulers, RCAR [6] have been developing exceptionally robust. The rules originating from spaces with mined categorical data. The RCAR algorithm forms classifiers in conjunction with the Regularized Cash Association Rules [8]. This algorithm progresses in three complicated steps: Firstly, mining a collection of Class Association Rules by utilizing predetermined support and confidence thresholds. The second step involves utilizing a regularized logistic regression technique with a Lasso penalty to develop a model that forecasts the conditional likelihood of an event happening. As Lasso regularization is selective, pruning useless rules occur in this step [9][10]. Lastly, by utilizing metarules and Lasso regularizations, the CARs are organized and visualized after making it through the initial stage of pruning. RCAR showcases itself as demonstrating high performance and accuracy in various fields, particularly in profiling and preventing security</p>
	<p>attacks [11]-[13].</p>

LONG	SHORT-TERM MEMORY(LSTM)	<p>LSM is stands of "Long Short- Term Memory". LMS units is use of learning recurrent neural networks "RNNs" those predicting the future from sequences at different lengths. Noticing that RNNs is workings with any kinds at se- quence data's, it is isn't limited to time series. LSMS is using a series of doors of control how information's enters, stores, or exits the net- works as sequences at data's. A typical LMS has four doors: a input gate, so has infor- mation's about cell states and a missing gate that delete information's that no longer neces- sary fo a model and a LSMT's output gate that selects the information's to presenting as a out- put. These gates can be thanked as filters, each forming they neural networks. This algorithm is widely using in time series forecasting as de- scribed by W. Bao, J. Yue & Y. Raso [2], M. Niou, G. Monsourfar & J.Bahezzadeh [7], or Taewook Kim & Ha Young Kim [14.]</p>
ARTIFICIALL	NEURAL NETWORK(ANN)	<p>A nerves network am a computer learning' sys- tem that utilize a function to interpret a certain type of input data 'n transform it into a specific output, commonly different from the input. An Artificial Neural Network (ANN) be approach to imitating' the interconnected network of neu- rons in the human brain, ending' computers to learn 'n make decisions akin to humans. Artifi- cial nerves networks function by programming' regular machines to mimic the intercon- nected brain cells, as demonstrated by (Wen Fenghua et al). Nerves networks excel in predictive analytics due to their hiding' layers. While the linear regression algorithm solely re- lies on input output nodes for prediction, nerves incorporate hiding' layers to enhance prediction accuracy. Essentially, they learn in a manner to humans!</p>

In the recent decades, technology advancements has been rapidly expanded and infiltrated every domain, including the stock market. Machine learning (ML in this sector involves the prediction of stock price, returns, and market trends. ML has revolutionized stock market practices, enabling investors to accurate

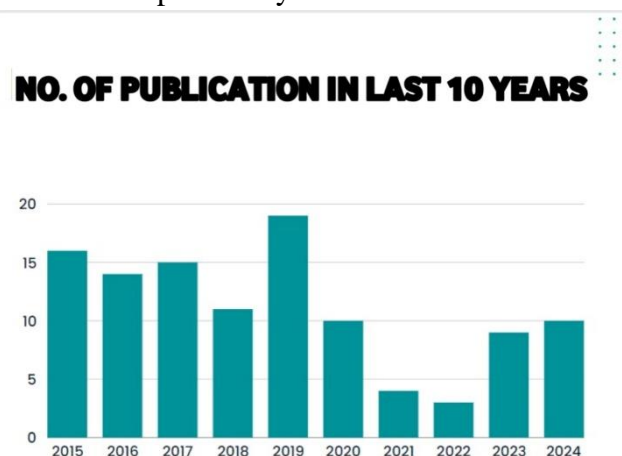
forecast stock market indices. Beyond price predictions, ML aids in understanding marketing behavior, offering innovating models for data processing and index predicting. This transformation has significantly altered the landscape of stock market forecasting. Various researchers have propose and testing new methodologies to enhancing the precision of stock market price predictions comparing to traditional techniques. Historically, stock market predictions relied theories like the efficient market hypothesis (EMH) and the random walk theory, alongside technical and fundamental analyses. Today, advancements in computing technology enable faster and more precis data processing, enhancing the efficacy of stock market predictions.

3. STOCK MARKET DATA

The study examine various stock market data classified for test with the aim to identify the most effective machine learn algorithm for stock market forecasting. In addition, the research focus on compare data from four stock market index: National Association of Securities Automatic Qu System (NASDAQ), York Stock Exchange (NYSE Nike and Financial Time Stock Exchange (FTSE All stock data spanning a decade from March 24, 2010, were sourced from Yahoo Finance. Yahoo Finance provides stock predesignate companies, where stock prices are influenced by economic, supply, and demand. Each market calculate differently; for instance, the FTSE index is base on the total market capitalization of the listed companies, while NASDAQ calculate base on the total value of stocks multiplied by closing price. NYSE calculation revolve around free-float market capitalization, and Nike's index is calculate by divide the sum of adjust price by a divisor. The datum set from Yahoo were download in an Excel file containing columns for stock prices, open and close prices, high and low price, enormous closing prices and trade volume. Moreover, NASDAQ and NYSE represent American indices, while FTSE and Nikkei represent British and Japanese indices respectively, each comprise specific set of company. For example, NASDAQ includes over 3,300 company, while FTSE encompass 100 company. The NYSE is the old index, following by NASDAQ and Nikkei, and the last among them is the FTSE, establish in 1984!![14]

4. COMPARATIVE ANALYSIS

In the lately years, they have seen an rise in the quantity of papers published. A peak of publications was saw in 2019, after a decreased in the two previous years.



Here in this bar graph, we can see the data of last ten years published paper count from 2015 to 2024. In that 2019 has the highest publication among the ten years. We can see from 2015 to 2019 have a comparatively higher number of publications done but from 2020 to 2024 we can see that there is some

gradual decrease in the publications. Even that there is only few publications in last five years but all of those published papers are trend setters.

5. CONCLUSION

In summary, machine learnings (MM) enhancers investors' ability to forecasting stock markets price trending more effectively than traditional methods caused by the acceleration and precisely data processed capabilities of computers. Popularities of MM applications for stock market fore castings reached peak in 2013, the period of 2015-2017 produced high numbers of related researches papers up to 15. Variety of machine learnings strategies are capable to deliver stock market forecasting with exceptionally accuracies, with some introducing innovations MM models to achieve even greater precision. While most MM models predominantly consist of neural networking (NNs), supported vector machines (SVM), and long short-term memories (LSTM) models are also utilized. It is hopes that these researches findings will AI-DS in expanding researchers' understandings of the most suitable approaches for predictions stock market movements.

The primary technique use for predicting stock market movements is neural networks with the LSTM algorithm being ' a derivative of this approach'. Neural networks account for 46.3% of the total methods employ, while SVM represents 17.1%, LSTM 19.5%, and other strategies compose the remaining 17.1%. Comparatively, the use of neural networks surpasses LSTM by 137.5%, SVM by 171.43%, and other models by the same percentage. LSTM gained popularity in 2015, whereas SVM saw more usage before 2016. In contrast, neural networks maintained consistent utilization over the span of ten years. Other methods like RF and KNN be less common in stock market forecasting.

Neural networks have the capacity to understand complex data trends, generalize ', an' adapt to the various relationships and patterns between inputs and outputs. This ability makes neural networks suitable for stock market prediction and results in more accurate forecasts compared to alternative methods. Consequently, neural networks remain the most prevalent model in the prediction of stock market trends.

6. REFERENCES

1. Kumar, I., Dogra, K., Utreja, C., & Yadav, P. (2018, April). A comparative study of supervised machine learning algorithms for stock market trend prediction. In 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT) (pp. 1003-1007). IEEE
2. Patel J, Shah S, Thakkar P, Kotecha K. Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. *Expert systems with applications*. 2015;42(1):259-68.
3. Vijh, M., Chandola, D., Tikkiwal, V. A., & Kumar, A. (2020). Stock closing price prediction using machine learning techniques. *Procedia computer science*, 167, 599- 606.
4. Fenghua, W. E. N., Jihong, X. I. A. O., Zhifang, H. E., & Xu, G. O. N.G. (2014). Stock price prediction based on SSA and SVM. *Procedia Computer Science*, 31, 625-631.
5. Vui CS, Soon GK, On CK, Alfred R, Anthony P. A review of stock market prediction with Artificial neural network (ANN). In: 2013 IEEE international conference on control system, computing and engineering. IEEE; 2013. p. 477- 82.
6. Azmi, M., Runger, G. C., & Berrado, A. (2019). Interpretable regularized class association rules algorithm for classification in a categorical data space. *Information Sciences*, 483,

313–331. <https://doi.org/10.1016/j.ins.2019.01.047>

7. Nikou, M., Mansourfar, G., & Bagherzadeh, J. (2019). Stock price prediction using DEEP learning algorithm and its comparison with machine learning algorithms. *Intelligent Systems in Accounting, Finance and Management*, 26(4), 164-174.
8. Towards an interpretable rules ensemble algorithm for classification in a categorical data space. In 2015 10th International Conference on Intelligent Systems: Theories and Applications (SITA) (pp. 1-6). IEEE.
9. Azmi, M., & Berrado, A. (2016, November). Class- association rules pruning using regularization. In 2016 IEEE/ACS 13th International Conference of Computer Systems and Applications (AICCSA) (pp. 1-7). IEEE.
10. Azmi, M., & Berrado, A. (2021). CARs-RP: Lasso- based class association rules pruning. *International Journal of Business Intelligence and Data Mining*, 18(2), 197-217.
11. Eddermoug, N., Mansour, A., Sadik, M., Sabir, E., & Azmi, M. (2021, June). KLM-based pro- filing and preventing security attacks for cloud computing: A comparative study. In 2021 28th International Conference on Telecommunications (ICT) (pp. 1-6). IEEE.
12. Eddermoug, N., Sadik, M., Sabir, E., Mansour, A., & Azmi, M. (2019, June). Ppsa: Profiling and preventing security attacks in cloud computing. In 2019 15th International Wireless Com- munications & Mobile Computing Conference (IWCMC) (pp. 415-421). IEEE.
13. Eddermoug, N., Mansour, A., Sadik, M., Sabir, E., & Azmi, M. (2019, October). Klm-PPSA: Klm- based profiling and preventing security attacks for cloud environments. In 2019 International Conference on Wireless Networks and Mobile Communications (WINCOM) (pp. 1-7). IEEE.
14. Vui CS, Soon GK, On CK, Alfred R, Anthony P. A review of stock market prediction with Ar- tificial neural network (ANN). In: 2013 IEEE international conference on control system, computing and engineering. IEEE; 2013. p. 477-82.
15. Najem, R., Meryem, F., Amr, A., & Talea, M. (Year). *Advancements in Artificial Intelligence and Machine Learning for Stock Market Prediction: A Comprehensive Analysis of Techniques and Case Studies*. *Procedia Computer Science* (pp.2-6)
16. Rouf, N., Malik, M. B., Arif, T., Sharma, S., Singh, S., Aich, S., & Kim, H.-C. (2021). *Stock Market Prediction Using Machine Learning Techniques: A Decade Survey on Methodologies, Recent Developments, and Future Directions*. *Electronics*.