International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 07 | July - 2022 Impact Factor: 7.185 ISSN: 2582-3930

A REVIEW ON PREDICTION OF STOCK MARKET USING MACHINE LEARNING

Sakshi Kesharwani

M.Tech

Shikha Singh

Assistant Professor

Saurabh Mitra

Head Of Department

Dept of Digital Communications, Dr. C.V.Raman University Kota, Bilaspur Chhattisgarh

Abstract: With the introduction of technical wonders such as worldwide digitalization, stock market forecasting has entered a technologically enhanced era, reviving the traditional trading methodology. Stock trading has become a center of investment for many financial investors as market capitalization continues to rise. Many analysts and researchers have created tools and approaches to forecast stock price fluctuations and assist investors in making informed decisions. Researchers can use advanced trading algorithms to forecast the market using non-traditional textual data from social media networks. The use of modern machine learning techniques like text data analytics and ensemble algorithms has substantially improved prediction accuracy. Meanwhile, due to its dynamic, irregular nature, stock market analysis and prediction remain one of the most difficult study disciplines. Based on the deployment of a generic framework, this paper illustrates the systematics of machine learning-based systems for stock market prediction. Findings from the previous decade (2011–2021) were evaluated using online digital libraries and databases such as the ACM digital library and Scopus. In addition, a thorough comparison study was performed to determine the direction of significance. The study will aid emerging researchers in understanding the fundamentals and accomplishments of this new field, allowing them to pursue additional research in promising paths.

Keywords: generic review; machine learning; stock market prediction; support vector machine

I. INTRODUCTION:

The equity capital markets serve as a marketplace for publicly traded corporations to issue and trade shares. Stocks, also known as shares, address fragmented ownership of an organization, resource, or security, and the stock market serves as a marketplace for investors to buy and sell responsibility for investable resources or offers [1]. Individuals that take part in such a stock and asset exchange are known as market participants. Domestic Retail Players, NRIs and Overseas Citizen of India (OCI)s, Domestic Institutions, Domestic Asset Management Companies (AMC) or Foreign Investors are the different types of market participants. Going public and issuing stocks that are then traded on the stock exchange is one option for businesses to raise capital for expansion or debt repayment n the secondary markets also known as stock exchanges. Instead of borrowing money in the form of cash, the company provides shares, which allows it to avoid losses, obligations, and interest payments. Second, to create money and earn profits for the stockholders. These shareholders, or investors, can earn [2] by investing in firms that pay regular dividends or by selling their shares on the market when the company's shares reach a greater rate than when they were acquired. Predicting the market value of stocks is therefore of tremendous importance to individuals who trade equities.

Stock Market Prediction refers to the effort to foresee or foretell the future value of a stock, a market sector, or even the whole market. It's an area that's piqued the interest of a wide range of people, including traders, market players, data analysts, and even computer engineers working in the fields of Machine Learning (ML) and Artificial Intelligence (AI), among others. Investing money in the stock market exposes you to a variety of market risks, as the value of a company's stock is highly dependent on its profits and performance in the marketplace, and can thus fluctuate due to a variety of factors such as government policies, microeconomic indicators, demand and supply, and so on.

These market variances are investigated in order to produce software and programmed utilizing various approaches such as machine learning, deep learning, neural networks, and artificial intelligence. Such systems and software can assist investors in appropriately anticipating the company's status based on historical and current data, market conditions, and other factors, and provide them with guidance in making decisions so that they do not lose their hard-earned money and maximize earnings.

The big data strategy, which tries to gain insights from a vast quantity of publicly accessible data and analyses this data on platforms such as Hadoop [3], is one of the ways for forecasting stock values. The deep learning technique is based on the use of neural networks to perform computations [4]. Long Short-Term Memory (LSTM) [5] is a form of Recurrent Neural Network (RNN) that is used to deal with long-term dependencies. Another technique to anticipate equity prices is to look into social media attitudes [6] or news items that may help determine the overall path that a company's or industry's shares may take based on a collective view. Because the value of a stock is sometimes viewed as a time series model, time series analysis [7] is a common methodology for predicting stock values.

We offer an overview of all the important technical methodologies used for stock price prediction in this study.

We looked at the numerous problems in each technique, as well as the future scope of each research, and made a comparison between them.

The quality of the characteristics used by stock market prediction algorithms has a significant impact on their success [9]. While several techniques for improving stock-explicit features have been employed by academics, feature extraction and selection require further attention mechanisms. The article's overview is shown in Figure 1.

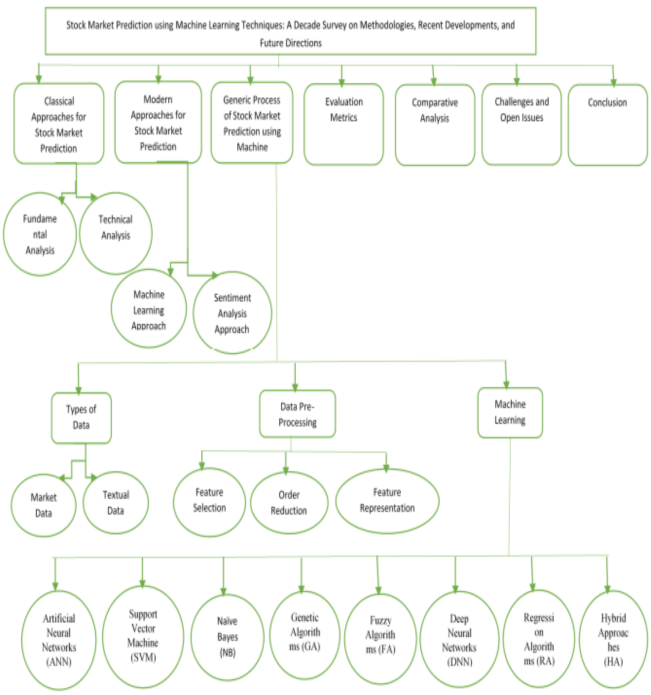

Fig 1 : Article Outline

II. LITERATURE REVIEW

In [1], Kim and Han developed a model for forecasting stock price index that combined artificial neural networks (ANN) and genetic algorithms (GAs) with feature discretization. The technical indicators as well as the direction of change in the daily Korea stock price index were employed in their research (KOSPI). Tey used data from 2928 trade days from January 1989 to December 1998 to determine their preferred features and formulae. Tey also used feature discretization optimization, which is an approach comparable to dimensionality reduction. The fact that they used GA to optimize the ANN is one of their work's merits. First, the buried layer has 12 input characteristics and processing components, not adjustable. Another disadvantage is that the authors only concentrated on two elements in the optimization phase throughout the ANN learning process. They still feel GA has a lot of promise when it comes to feature discretization optimization. Our initialized feature pool refers to the features that have been chosen.

In [2], Qiu and Song provided a method based on an optimized artificial neural network model to forecast the direction of the Japanese stock market. The authors use genetic algorithms in conjunction with artificial neural network-based models to create a hybrid GA-ANN model.

In [3], Piramuthu evaluated several feature selection strategies for data mining applications in depth. He evaluated how different feature selection strategies maximized decision tree performance using datasets such as credit approval data, loan default data, online traffic data, tam, and kiang data. The Bhattacharyya measure, the Matusita measure, the divergence measure, the Mahalanobis distance measure, and the Patrick-Fisher measure were among the probabilistic distance measures he compared. The Minkowski distance measure, city block distance measure, Euclidean distance measure, Chebychev distance measure, and nonlinear (Parzen and hyper-spherical kernel) distance measure are all inter-class distance measures. The author assessed both probabilistic distance-based and multiple inter-class feature selection approaches, which is a strength of this study. Furthermore, the author conducted the study using a variety of datasets, which added to the paper's strength. The evaluation algorithm, on the other hand, was only a decision tree. We can't say if the feature selection approaches would hold up in a bigger dataset or with a more complicated model.

In [4], Hassan and Nath used the Hidden Markov Model (HMM) to estimate stock market values for four major airlines. They divide the model's states into four categories: opening price, closing price, maximum price, and lowest price. The technique used in this study is unique in that it does not require expert knowledge to develop a prediction model. While this research is confined to the airline sector and assessed on a short dataset, it may not result in a generalizable prediction model. To accomplish the comparison job, one of the methodologies used in stock market prediction related activities might be used. The date range of the training and testing datasets was specified at a maximum of two years by the authors which provided us a date range reference for our evaluation part.

Wavelet Neural Network (WNN) was used by **Lei in [5]** to anticipate stock price patterns. As an optimization, the author used Rough Set (RS) for attribute reduction. The stock price trend feature dimensions were reduced using Rough Set. It was also used to identify the Wavelet Neural Network's structure. This study's dataset includes five well-known stock market indices: (1) SSE Composite Index (China), (2) CSI 300 Index (China), (3) All Ordinaries Index (Australia), (4) Nikkei 225 Index

(Japan), and (5) Dow Jones Index (USA). The model was evaluated using several stock market indexes, and the results were convincing in terms of generality. The computational complexity is reduced by employing Rough Set to optimize the feature dimension before processing. However, in the discussion section, the author mainly emphasized parameter tweaking and did not mention the model's flaws. Meanwhile, we discovered that while the assessments were done on indices, the same model would not perform as well if applied to an individual stock.

Lee in [6] employed the support vector machine (SVM) in conjunction with a hybrid feature selection strategy to forecast stock movements. In this study, the dataset is a subset of the NASDAQ Index from the Taiwan Economic Journal Database (TEJD) in 2008. The wrapper for the feature selection component was a hybrid technique that used supported sequential forward search (SSFS). Another benefit of this study is that it included a comprehensive approach for parameter tuning that included performance under various parameter values. The feature selection model's obvious structure is also a heuristic for the initial stage of model structuring. One of the drawbacks was that the performance of SVM was only compared to that of a back-propagation neural network (BPNN) and not to that of other neural networks the other machine learning algorithms.

Sirignano and Cont in [7] used a deep learning solution that was trained on a universal collection of financial market features the dataset comprised purchase and sell records for all transactions, as well as cancellations of orders for about 1000 NASDAQ equities through the stock exchange's order book. Te NN is made up of three layers, each with LSTM units, followed by a feed-forward layer with rectified linear units (ReLUs) and the stochastic gradient descent (SGD) method as an optimization. Teir universal model was able to generalize and cover stocks not in the training data. Even though they acknowledged the benefits of a universal paradigm, the cost of training was still high. Meanwhile, it's unclear whether there are any unnecessary characteristics due to the deep learning algorithm's inexplicit programming contaminated when feeding the data into the model. Authors found out that it would have been better if they performed feature selection part before training the model and found it as an effective way to reduce the computational complexity.

Ni et al. in [8] predicted SVM was used to analyses stock price trends, and fractal feature selection was used for optimization. The Shanghai Stock Exchange Composite Index (SSECI) was utilized as the dataset, which included 19 technical indicators. They optimized the input data before processing it by doing feature selection. They also utilized a grid search approach called k cross-validation to discover the optimum parameter combination. Furthermore, the examination of various feature selection approaches is extensive. According to the authors' conclusion, they only looked at technical indications in the financial arena and ignored macro and micro issues. The authors' dataset source was similar to our dataset; hence their assessment results are applicable to our research. Tey also discussed the k technique.

McNally et al. in [8] RNN and LSTM were used to forecast the price of Bitcoin, and the feature engineering aspect was improved using the Boruta method, which works similarly to the random forest classifier. They employed Bayesian optimization to select LSTM parameters in addition to feature selection. The Bitcoin data was collected between August 19th, 2013 and July 19th, 2016. To increase the performance of deep learning algorithms, different optimization strategies were used. Overfitting is the main issue in their research. The issue of anticipating Bitcoin price trends is comparable to that of predicting stock market prices. This work is threatened by hidden characteristics and noise buried in the pricing data. The study subject was approached as a time sequence problem by the writers. The most interesting portion of this study is the feature engineering and optimization part; we could replicate the methods they exploited in our data pre-processing.

Weng et al. In [9] targeted on quick-time period stock rate prediction by using the use of ensemble methods of 4 famous machine mastering models. Te dataset for this studies is fve units of statistics. Tey acquired these datasets from 3 open-sourced APIs and an R package named TTR. Te gadget learning models they was (1) neural community regression ensemble (NNRE), (2) a Random Forest with unpruned regression trees as base newbies (RFR), (three) AdaBoost with unpruned regression bushes as base newbies (BRT) and (four) a aid vector regression ensemble (SVRE). A thorough take a look at of ensemble techniques precise for quick-term stock price prediction. With historical past understanding, the authors selected eight technical indicators in this have a look at then done a considerate assessment of five datasets. Te primary contribution of this paper is they developed a platform for investors the usage of R, which does now not want customers to input their personal statistics but call API to fetch the facts from online supply honest. From the research angle, they only evaluated the prediction of the rate for 1 up to ten days beforehand however did no longer compare longer terms than two buying and selling weeks or a shorter time period than 1 day. Te number one dilemma in their studies changed into that they simplest analyzed 20 U.S.-based totally shares, the version might not be generalized to different inventory market or want further revalidation to look if it suffered from overfitting problems.

Kara et al. In [10] also exploited ANN and SVM in predicting the motion of inventory rate index. Te facts set they used covers a time period from January 2, 1997, to December 31, 2007, of the Istanbul Stock Exchange. Te primary power of this work is its specified report of parameter adjustment procedures. While the weaknesses of this work are that neither the technical indicator nor the version structure has novelty, and the authors did not provide an explanation for how their version executed better than different fashions in preceding works. Tus, greater validation works on different datasets would help. Tey defined how ANN and SVM work with inventory market functions, also recorded the parameter adjustment. Te implementation part of our research should gain from this preceding work.

Jeon et al. In [11] completed studies on millisecond c program language period-primarily based big dataset by using the use of pattern graph monitoring to complete stock fee prediction tasks. Te dataset they used is a millisecond c language-based totally big dataset of ancient inventory information from KOSCOM, from August 2014 to October 2014, 10G–15G capability. Te author carried out Euclidean distance, Dynamic Time Warping (DTW) for sample reputation. For characteristic selection, they used stepwise regression. Te authors completed the prediction mission by way of ANN and Hadoop and RHive for massive information processing. Te “Results” segment is based totally on the result processed through a combination of SAX and Jaro–Winkler distance. Before processing the records, they generated aggregated statistics at 5-min intervals from discrete data. Te primary power of this paintings is the express shape of the whole implementation system. While they exploited a fantastically antique version, some other weak point is the general time span of the schooling dataset is extraordinarily short. It is tough to get admission to the millisecond c language-based records in real existence, so the model is not as sensible as a daily primarily based data model.

Huang et al. In [12] carried out a fuzzy-GA version to finish the stock choice project. Tey used the key shares of the 2 hundred biggest marketplace capitalization listed as the funding universe within the Taiwan Stock Exchange. Besides, the yearly monetary assertion facts and the inventory returns were taken from the Taiwan Economic Journal (TEJ) database at www. Tej.Com.Tw/ for the term from 12 months 1995 to 12 months 2009. Tey performed the bushy membership feature with model parameters optimized with GA and extracted features for optimizing inventory scoring. Te authors proposed an optimized version for selection and scoring of stocks. Different from the prediction model, the authors more focused on inventory rankings, choice, and performance evaluation. Teir shape is extra sensible among traders. But inside the version validation component, they did no longer examine the version with existed algorithms but the statistics of the benchmark, which made it tough to pick out if GA could outperform different algorithms.

Fischer and Krauss in [13] applied long short-time period memory (LSTM) on monetary market prediction. Te dataset they used is S&P 500 index constituents from Tomson Reuters. Tey obtained all month-cease constituent lists for the S&P 500 from Dec 1989 to Sep 2015, then consolidated the lists into a binary matrix to eliminate survivor bias. Te authors also used RMS prop as an optimizer, that is a mini-batch version of rprop. Te primary power of this work is that the authors used the today's deep getting to know approach to carry out predictions. Tey depended on the LSTM method, lack of history know-how in the financial domain. Although the LSTM outperformed the usual DNN and logistic regression algorithms, at the same time as the writer did not mention the attempt to train an LSTM with long-time dependencies.

Tsai and Hsiao in [14] proposed an answer as a mixture of various characteristic choice methods for prediction of shares. Tey used Taiwan Economic Journal (TEJ) database as statistics supply. Te statistics used of their analysis become from yr 2000 to 2007. In their paintings, they used a sliding window technique and combined it with multi-layer perceptron (MLP) primarily based synthetic neural networks with lower back propagation, as their prediction version. In their work, in addition they applied major issue analysis (PCA) for dimensionality reduction, genetic algorithms (GA) and the type and regression timber (CART) to pick essential features. Tey did now not simply rely upon technical indices only. Instead, additionally they included each essential and macroeconomic indices of their evaluation. Te authors also said a comparison on characteristic selection strategies. Te validation component became accomplished via combining the version performance stats with statistical evaluation.

Pimenta et al. In [15] leveraged an automatic investing technique by way of the use of multi-objective genetic programming and implemented it inside the inventory market. Te dataset was acquired from Brazilian stock trade marketplace (BOVESPA), and the primary strategies they exploited were a combination of multi-objective optimization, genetic programming, and technical buying and selling policies. For optimization, they leveraged genetic programming (GP) to optimize decision guidelines. Te novelty of this paper become within the evaluation component. Tey included a ancient duration, which changed into a critical second of Brazilian politics and economics whilst performing validation. Tis approach reinforced the generalization energy in their proposed version. When deciding on the sub-dataset for assessment, in addition they set criteria to ensure greater asset liquidity. While the baseline of the contrast became too fundamental and essential, and the authors did now not perform any evaluation with different present fashions.

Huang and Tsai in [16] performed a filter out-based feature choice assembled with a hybrid self-organizing feature map (SOFM) help vector regression (SVR) model to forecast Taiwan index futures (FITX) trend. Tey divided the education samples into clusters to marginally improve the schooling performance. Te authors proposed a comprehensive version, which became a mixture of two novel gadget getting to know techniques in stock marketplace analysis. Besides, the optimizer of characteristic choice become additionally applied before the statistics processing to improve the prediction accuracy and decrease the computational complexity of processing every day inventory index records. Tough they optimized the function selection element and split the pattern information into small clusters, it was already strenuous to educate each day inventory index facts of this model. It might be hard for this version to expect trading sports in shorter time durations for the reason that facts volume could be increased considerably. Moreover, the evaluation is not strong enough in view that they set a single SVR model as a baseline, however did no longer examine the performance with other preceding works, which brought on trouble for future researchers to identify the advantages of SOFM-SVR model why it outperforms other algorithms.

Takur and Kumar in [17] also developed a hybrid economic trading help gadget by exploiting multi-class classifiers and random wooded area (RAF). Tey conducted their research on inventory indices from NASDAQ, DOW JONES, S&P 500, NIFTY 50, and NIFTY BANK. Te authors proposed a hybrid version combined random forest (RF) algorithms with a weighted multicategory generalized eigenvalue help vector machine (WMGEPSVM) to generate "Buy/Hold/Sell" signals. Before processing the records, they used Random Forest (RF) for feature pruning. Te authors proposed a sensible model designed for actual-existence funding activities, that can generate three primary indicators for traders to consult. Tey also performed a thorough contrast of related algorithms. While they did now not mention the time and computational complexity of their works. Meanwhile, the unignorable problem of their work was the lack of economic area information background. Te traders regard the indices records as one of the attributes but couldn't take the signal from indices to operate a specific inventory truthful.

Hsu in [18] assembled characteristic selection with a again propagation neural network (BNN) blended with genetic programming to predict the stock/futures price. Te dataset on this research turned into received from Taiwan Stock Exchange Corporation (TWSE). Te authors have brought the outline of the heritage understanding in element. While the weak spot in their work is that it is a lack of records set description. This is an aggregate of the version proposed by means of other preceding works. Tough we did now not see the newness of this work, we will nonetheless conclude that the genetic programming (GP) set of rules is admitted in stock marketplace studies area. To reinforce the validation strengths, it would be right to don't forget including GP fashions into assessment if the model is predicting a specific rate.

Hafezi et al. In [19] constructed a bat-neural community multi-agent machine (BN-NMAS) to predict inventory rate. Te dataset was received from the Deutsche bundes-financial institution. Tey also implemented the Bat algorithm (BA) for optimizing neural community weights. Te authors illustrated their standard structure and good judgment of device layout in clean flowcharts. While there had been only a few preceding works that had performed on DAX information, it would be difficult to recognize if the model they proposed still has the generality if migrated on other datasets. Te gadget design and characteristic choice logic are captivating, which really worth relating to. Teir findings in optimization algorithms are also treasured for the studies in the stock market rate prediction studies area. It is worth trying the Bat set of rules (BA) whilst building neural community fashions.

Long et al. In [20] carried out a deep learning method to expect the stock rate motion. Te dataset they used is the Chinese inventory market index CSI 300. For predicting the stock price motion, they built a multi-flter neural network (MFNN) with stochastic gradient descent (SGD) and again propagation optimizer for getting to know NN parameters. Te energy of this paper is that the authors exploited a novel model with a hybrid version constructed through distinct sorts of neural networks, it provides an notion for building hybrid neural community structures.

Atsalakis and Valavanis in [21] proposed an answer of a neuro-fuzzy system, which consists of controller named as Adaptive Neuro Fuzzy Inference System (ANFIS), to achieve brief-time period inventory charge trend prediction. Te major electricity of this work is the assessment element. Not simplest did they examine their proposed gadget with the popular statistics fashions, however also compared with investment strategies. While the weak point that we discovered from their proposed answer is that their solution architecture is lack of optimization element, which might limit their model overall performance. Since our proposed answer is likewise that specialize in quick-term stock price fashion prediction, this work is heuristic for our gadget layout. Meanwhile, by using comparing with the popular buying and selling techniques from traders, their work inspired us to evaluate the strategies utilized by buyers with techniques utilized by researchers.

Nekoeiqachkanloo et al. In [22] proposed a machine with two one of a kind method for inventory investment. Te strengths of their proposed answer are obvious. First, it's far a comprehensive device that consists of records pre-processing and two exceptional algorithms to indicate the nice investment portions. Second, the system also embedded with a forecasting component, which also keeps the functions of the time collection. Last but not least, their input functions are a mixture of essential features and technical indices that goal to fll within the hole among the economic domain and technical domain. However, their paintings have a weakness within the assessment part. Instead of comparing the proposed machine on a big dataset, they chose 25 famous shares. Tere is a high opportunity that the well-known stocks may doubtlessly share some common hidden functions.

As some other associated contemporary paintings, **Idrees et al. [23]** published a time series-based prediction approach for the volatility of the inventory marketplace. ARIMA is not a new approach in the time collection prediction research domain. Teir work is extra focusing at the characteristic engineering aspect. Before feeding the capabilities into ARIMA models, they designed three steps for function engineering: Analyze the time series, identify if the time series is stationary or not, carry out estimation by way of plot ACF and PACF charts and search for parameters. Te simplest weak point of their proposed solution is that the authors did not carry out any customization on the existing ARIMA version, which would possibly limit the gadget overall performance to be stepped forward.

III. PROPOSED SYSTEM:

We use machine learning methods like Random Forest and Support Vector Machines to estimate stock prices in this suggested system. We presented the "Stock market price prediction" system, and we used the random forest algorithm to forecast the stock market price. We were able to train the computer using multiple data points from the past to produce a future forecast in our suggested method. To train the algorithm, we used data from the previous year's stocks. To tackle the challenge, we primarily employed two machine-learning packages. The first was numpy, which was used to clean and alter the data before putting it in a format that could be analyzed. Scikit, on the other hand, was employed for real-time analysis and prediction. We used data from past years' stock markets, which we obtained from a public database available online; 80% of the data was used to train the machine, while the remaining 20% was used to test the data. The supervised learning model's core strategy is to learn patterns and correlations in data from the training set and then replicate them for the test data. For data processing, we utilized the Python pandas module, which merged many datasets into a single data frame. We were able to prepare the data for feature extraction using the fine-tuned data frame. The date and the closing price for a certain day were the data frame’s characteristics. All of these characteristics were utilized to train the computer using a random forest model to predict the object variable, which is the price for a specific day. We also calculated the accuracy by comparing the test set predictions to the actual values. Data pre-processing, random forest, and other fields of research are all touched on by the suggested approach.

IV. CONCLUSION:

We looked at how multiple global financial markets may be used with machine learning algorithms to anticipate stock index changes in this research. The SVM technique using machine learning is used to analyses a vast collection of data collected from several worldwide financial marketplaces. Furthermore, SVM does not have the issue of overfitting. For predicting the daily trend of Market stocks, many machine learning-based algorithms have been presented. The remarkable efficiency is supported by numerical data. Our well-trained predictor was used to create realistic trading models. In comparison to the chosen benchmarks, the model yields a bigger profit.

REFERENCE:

[1]. Kim K, Han I. Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index. Expert Syst Appl. 2000;19:125–32. https://doi.org/10.1016/S0957-4174(00)00027-0

[2]. Qiu M, Song Y. Predicting the direction of stock market index movement using an optimized artificial neural net-work model. PLoS ONE. 2016;11(5):e0155133.

[3]. Piramuthu S. Evaluating feature selection methods for learning in data mining applications. Eur J Oper Res. 2004;156(2):483–94. https://doi.org/10.1016/S0377-2217(02)00911-6.

[4]. Hassan MR, Nath B. Stock market forecasting using Hidden Markov Model: a new approach. In: Proceedings—5th international conference on intelligent systems design and applications 2005, ISDA’05. 2005. pp. 192–6. https://doi.org/10.1109/ISDA.2005.85.

[5]. Lei L. Wavelet neural network prediction method of stock price trend based on rough set attribute reduction. Appl Soft Comput J. 2018;62:923–32. https://doi.org/10.1016/j.asoc.2017.09.029.

[6]. Lee MC. Using support vector machine with a hybrid

feature selection method to the stock trend prediction. Expert Syst Appl. 2009;36(8):10896–904. https://doi.org/10.1016/j.eswa.2009.02.038.

[7]. Sirignano J, Cont R. Universal features of price formation in financial markets: perspectives from deep learning. Ssrn. 2018. https://doi.org/10.2139/ssrn.3141294.

[8]. Ni LP, Ni ZW, Gao YZ. Stock trend prediction based on fractal feature selection and support vector machine. Expert Syst Appl. 2011;38(5):5569–76. https://doi.org/10.1016/j.eswa.2010.10.079.

[9]. McNally S, Roche J, Caton S. Predicting the price of bitcoin using machine learning. In: Proceedings—26th Euromi- cro international conference on parallel, distributed, and network-based processing, PDP 2018. pp. 339–43. https:// doi.org/10.1109/PDP2018.2018.00060.

[10].Weng B, Lu L, Wang X, Megahed FM, Martinez W. Predicting short-term stock prices using ensemble methods and online data sources. Expert Syst Appl. 2018;112:258–73. https://doi.org/10.1016/j.eswa.2018.06.016.

[11]. Kara Y, Acar Boyacioglu M, Baykan ÖK. Predicting direction of stock price index movement using artificial neural net-works and support vector machines: the sample of the Istanbul Stock Exchange. Expert Syst Appl. 2011;38(5):5311– 9. https://doi.org/10.1016/j.eswa.2010.10.027.

[12]. Jeon S, Hong B, Chang V. Pattern graph tracking-based stock price prediction using big data. Future Gener Comput Syst. 2018;80:171–87. https://doi.org/10.1016/j.future.2017.02.010.

[13]. Huang CF, Chang BR, Cheng DW, Chang CH. Feature selection and parameter optimization of a fuzzy-based stockselection model using genetic algorithms. Int J Fuzzy Syst. 2012;14(1):65–75. https://doi.org/10.1016/J.POLYM ER.2016.08.021.

[14].Fischer T, Krauss C. Deep learning with long short-term memory networks for financial market predictions. Eur J Oper Res. 2018;270(2):654–69. https://doi.org/10.1016/j.ejor.2017.11.054.

[15].Tsai CF, Hsiao YC. Combining multiple feature selection methods for stock prediction: union, intersection, and multi- intersection approaches. Decis Support Syst. 2010;50(1):258–69. https://doi.org/10.1016/j.dss.2010.08.028.

[16]. Pimenta A, Nametala CAL, Guimarães FG, Carrano EG. An automated investing method for stock market based on multiobjective genetic programming. Comput Econ. 2018;52(1):125–44. https://doi.org/10.1007/s10614-017-9665-9.

[17].Huang CL, Tsai CY. A hybrid SOFM-SVR with a filter-based feature selection for stock market forecasting. Expert Syst Appl. 2009;36(2 PART 1):1529–39. https://doi.org/10.1016/j.eswa.2007.11.062.

[18].Thakur M, Kumar D. A hybrid financial trading support system using multi-category classifiers and random forest. Appl Soft Comput J. 2018;67:337–49. https://doi.org/10.1016/j.asoc.2018.03.006.

[19]. Hsu CM. A hybrid procedure with feature selection for resolving stock/futures price forecasting problems. Neural Comput Appl. 2013;22(3–4):651–71. https://doi.org/10.1007/s00521-011-0721-4.

[20].Hafezi R, Shahrabi J, Hadavandi E. A bat-neural network multi-agent system (BNNMAS) for stock price prediction: case study of DAX stock price. Appl Soft Comput J. 2015;29:196–210. https://doi.org/10.1016/j.asoc.2014.12.028.

[21]. Long W, Lu Z, Cui L. Deep learning-based feature engineering for stock price movement prediction. Knowl Based Syst. 2018;164:163–73. https://doi.org/10.1016/j.knosys.2018.10.034.

[22]. Atsalakis GS, Valavanis KP. Forecasting stock market short-term trends using a neuro-fuzzy based methodology. Expert Syst Appl. 2009;36(7):10696–707.

[23]. Nekoeiqachkanloo H, Ghojogh B, Pasand AS, Crowley M. Artificial counselor system for stock investment. 2019. ArXivPreprint arXiv:1903.00955.

[24]. Idrees SM, Alam MA, Agarwal P. A prediction approach for stock market volatility based on time series data. IEEE Access. 2019;7:17287–98. https://doi.org/10.1109/ACCESS.2019.2895252