Computer Science Review 34 (2019) 100190

Contents lists available at ScienceDirect

Computer Science Review

journal homepage: www.elsevier.com/locate/cosrev

# Systematic analysis and review of stock market prediction techniques

Dattatray P. Gandhmal, K. Kumar

*School of Computer Science & Engineering, Vellore Institute of Technology, Vellore, India*

ARTICLE INFO

*Article history:*
Received 3 April 2019
Received in revised form 2 July 2019
Accepted 6 August 2019
Available online 28 August 2019

*Keywords:*
Stock market prediction
Bayesian model
Fuzzy classifier
ANN
Classification
Clustering

ABSTRACT

Prediction of stock market trends is considered as an important task and is of great attention as predicting stock prices successfully may lead to attractive profits by making proper decisions. Stock market prediction is a major challenge owing to non-stationary, blaring, and chaotic data, and thus, the prediction becomes challenging among the investors to invest the money for making profits. Several techniques are devised in the existing techniques to predict the stock market trends. This work presents the detailed review of 50 research papers suggesting the methodologies, like Bayesian model, Fuzzy classifier, Artificial Neural Networks (ANN), Support Vector Machine (SVM) classifier, Neural Network (NN), Machine Learning Methods and so on, based on stock market prediction. The obtained papers are classified based on different prediction and clustering techniques. The research gaps and the challenges faced by the existing techniques are listed and elaborated, which help the researchers to upgrade the future works. The works are analyzed using certain datasets, software tools, performance evaluation measures, prediction techniques utilized, and performance attained by different techniques. The commonly used technique for attaining effective stock market prediction is ANN and the fuzzy-based technique. Even though a lot of research efforts, the current stock market prediction technique still have many limits. From this survey, it can be concluded that the stock market prediction is a very complex task, and different factors should be considered for predicting the future of the market more accurately and efficiently.



**Contents**



## 1. Introduction

The advancements in stock price prediction have gained significant importance among expert analysts and investors. The stock market prediction for analyzing the trends is complicated due to intrinsic noisy environments and large volatility with respect to

*E-mail addresses:* gdattatray@gmail.com (D.P. Gandhmal), kkumar@vit.ac.in (K. Kumar).

https://doi.org/10.1016/j.cosrev.2019.08.001
1574-0137/

the market trends. The complexities of the stock prices adapt certain factors that involve quarterly earnings' reports, market news, and varying changing behaviors. The traders depend on various technical indicators that are based on the stocks, which are collected on a daily basis. Even though these indicators are used to analyze the stock returns, it is complicated to forecast daily and weekly trends in the market [1]. The accurate prediction of stock trends is interesting and a complex task in the changing industrial world. Several aspects, which affect the behavior of stock trends, are non-economic and economic factors and which are taken into consideration. Thus, predicting the stock market is considered as a major challenge for increasing production [2]. Traditional techniques reveal that the stock market earnings are predicted from previous stock returns and other financial variables and macroeconomics. The prediction of stock market revenues directed the investors towards examining the causes of predictability. The forecasting of stock trends is a difficult process as it is influenced by several aspects, which involve trader's expectations, financial circumstances, administrative events, and certain aspects related to the market trends. Moreover, the list of stock prices is usually dynamic, complicated, noisy, nonparametric, and nonlinear by nature [3]. The forecasting of financial time series becomes an issue due to certain complex features, like volatility, irregularities, noise, and changing trends [4].

Various models applied for predicting the stock prices are managed using the time series models that involve Auto-Regressive Conditional Heteroscedastic (ARCH) model, Generalized Auto-Regressive Moving Average (GARCH), and Auto-Regressive Moving Average (ARMA). However, these models entail historical data and hypothesis like normality postulates. Several methods used for stock market prediction are based on conventional time series, such as fuzzy time series data, real numbers, and design of fuzzy sets. The fuzzy time series data are implemented for stock market prediction for handling linguistic value data for producing precise predicting results. These methods are widely used for forecasting nonlinear and dynamic datasets in the changing domains, such as tourism demand and stock markets [5]. Many intelligent techniques, namely soft computing algorithms, Neural Network (NN), back propagation algorithm, and Genetic Algorithm (GA), are applied for predicting the stock market returns. In [6], a prediction model was designed for predicting the stock trends with time series models. In [7], GA and NN are integrated for designing hybrid expert systems to make the investment decisions. A technique based on GA is designed in [8], for feature discretization and determining the weights of Artificial Neural Networks (ANNs) [9] for predicting the index of the stock price. Time series model and NN were combined to predict the variability of the stock price in [10]. The Artificial Intelligence (AI) techniques, like ANN, were devised for predicting the stock market prices. Many networks used feed forward neural networks for predicting the stock trends and evaluated multiple parametric and non-parametric models to forecast the stock market returns [11]. Soft computing methods are utilized to deal with the AI for making the decisions using the profit and loss criterions. The techniques employed are fuzzy logic [12], Particle Swarm Optimization (PSO) [13], ANN [14,15], and Support Vector Machine (SVM) [16,17]. Several researchers tried to employ fuzzy based techniques and randomness for optimizing the pricing models [18,19]. In [20], the fuzzy-based techniques are employed for analyzing the market trends, and in [21], the performance of the fuzzy forecast is derived for estimating the initial values of stock price [22].

The primary intention of this work is to give a detailed review using various stock market prediction methodologies for predicting future trends and stock returns. This survey considers several existing stock market prediction techniques that are adapted in the research works. The review is made by taking the adapted methodologies, publication year, performance evaluation metrics, datasets employed, and software tools. In addition, the performance attained by different techniques is used for predicting the stock market trends. Here, the existing methods of stock market prediction have been categorized into distinct classification and clustering techniques, and further, the survey is performed for the exploitation of the research gap and issues found in those techniques. Thus, the survey acts as the motivation for the future extension of effective stock market prediction.

The paper is structured in the following manner: Section 1 presents a brief introduction of stock trend prediction, Section 2 provides needs and description regarding the prediction of the stock market, and literature survey of the existing stock prediction techniques and Section 3 elucidates analysis and discussion of the existing techniques. Section 4 deliberates the research gaps and issues, and the future works and Section 5 provides the conclusion of the survey.

## 2. Stock market prediction: Need and description

The stock market has gained the attraction of the investors due to advanced applications, in which the forecasting may lead to successful market prediction. The prediction of the stock trends directly depends on investing and trading of stock data. The tools employed for the stock market prediction can monitor, predict, and regulate the market, which can be utilized for taking correct decisions [23]. The stock market has to deal with multiple industrial stock data, which covers the whole financial market. With respect to the market status, the actions are adapted by the investors considering the sales and purchase. Several factors, which influence the status of the market, are estimations of future incomes, a news release on profits, declaration of dividends, management changes, and so on. The research in stock trading issues led to the prediction of some aspects, which are considered by the experts for influencing the stock price [24]. The prediction mechanisms based on the stock market play an important role to bring more persons and existing investors in common place. The accurate results for predicting the stock market help the investors for making better decisions. The data mining tools can help the investors to predict future trends and behaviors and help the institutions for yielding active solutions to make knowledge-driven decisions [25]. The intelligent data analysis helps the tools to generate results with ease. The extraction of useful information is an effective way for mining data [26]. Several data mining techniques [27] and knowledge discovery from databases [28,29] have been employed for analyzing the market trends. Here, the data mining is essential for predicting the stock market, which can search the hidden parts and increase the accuracy levels for analyzing the market trends using techniques, such as regression approach, Knowledge Discovery in Databases (KDD), fuzzy models for making effective investment decisions [30]. Fig. 1 depicts the functioning of the stock market prediction system.

Fig. 1 illustrates the block diagram of the function of the stock market prediction system. Initially, the historical stock data are collected from various datasets, such as Bombay Stock Exchange (BSE) dataset, Taiwan Stock Exchange Weighted Index (TAIEX), and so on, based on a technical index. Here, the technical index is defined as the relevance of the product to be utilized in a specific application. Then, the collected data is subjected to data preprocessing, for removing the noise and artifacts from the datasets. At first, the data needs attribute relevance analysis to be applied for removing the unwanted attributes. Then, the pre-processed data are employed for selecting the significant features, which

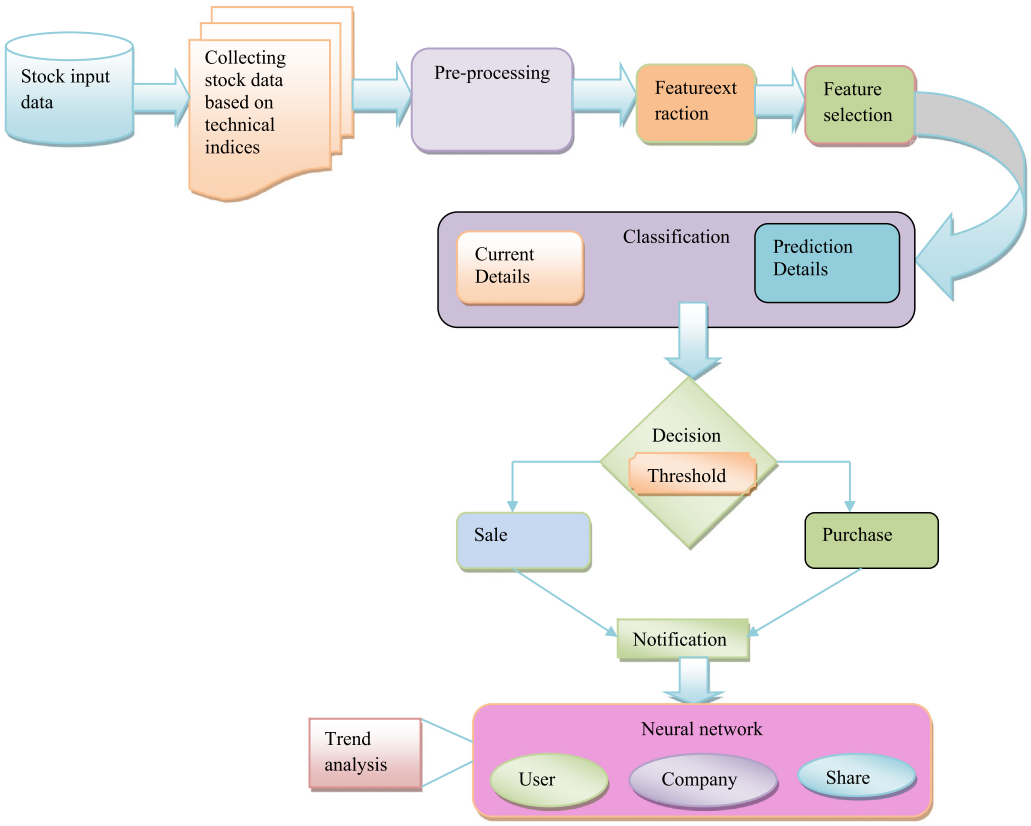


**Fig. 1.** Block diagram of the functioning of the stock market prediction system.

can be useful for predicting the stock trends. Then, the selected feature is analyzed for extracting the knowledge using a large amount of data. The data analyzer provides improved analysis and user-friendly interface [31]. Here, the data analyzer divides the obtained details into two classes, namely prediction details and current details, which are considered as a foundation for decision making. The prediction details, the current details, and the threshold values are subjected to the decision-making system. Using the threshold value, the pre-processing is done to declare profit or loss. Then, the alert notification is generated based on the status. If the investor is in profit, then the shares can be used for increasing the sale, and if in loss, then this share is given more attention for its development, and the NN [32] is applied to make effective decisions.

### *2.1. Literature survey on different stock market prediction techniques*

This section describes the review of different stock market prediction techniques. Fig. 2 illustrates the categorization of distinct stock market prediction techniques. The stock market prediction techniques are broadly categorized into two types, namely prediction based techniques and clustering based techniques. The techniques based on ANN, Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Decision Support System (DSS), Hidden Markov Model (HMM), Naive Bayes (NB), NN, Support Vector Regression (SVR) and SVM are grouped under prediction based techniques. Likewise, the techniques based on filtering, fuzzy, k-means, and optimization are grouped under clustering based techniques.

### *2.2. Classification schemes for stock market prediction*

The specific research works that employ different classification schemes are elaborated in this section.

**(1) Analysis based on prediction techniques**

The analysis based on prediction techniques is elaborated in this section, which includes ANN, CNN, DSS, HMM, NN, RNN, SVM, SVR, and NB.

**(a) ANN-based prediction techniques**

ANN captures the structural relationship between a stock's performance and its determinant factors more accurately than many other statistical methods. In literature, various sets of input variables are utilized to predict stock returns. Some researchers pre-processed the input variables before applied it to the ANN for prediction. This subsection elaborates different research works, which employed the ANN-based stock market prediction as follows,

Ticknor, J.L [1] designed a model named Bayesian regularized ANN, for predicting the behavior of the financial market. Market trends and technical indicators are utilized as inputs for predicting the cost of each stock. Accurate forecasting of stock prices is important to serve the investors to enhance the stock returns. Here, the Bayesian regularized network allocates weight to the permitted the network for dealing with complicated models. The model decreases the potential to improve the quality of prediction and network generalization, but it failed to consider technical indicators, which enhances the model quality for upcoming applications. Rout, A.K et al. [33] developed Computationally Efficient Functional Link Artificial Neural Network (CEFLANN) on the basis of an adaptive model for predicting the financial time series data using Indian stock market indexes. The model adopts the Least

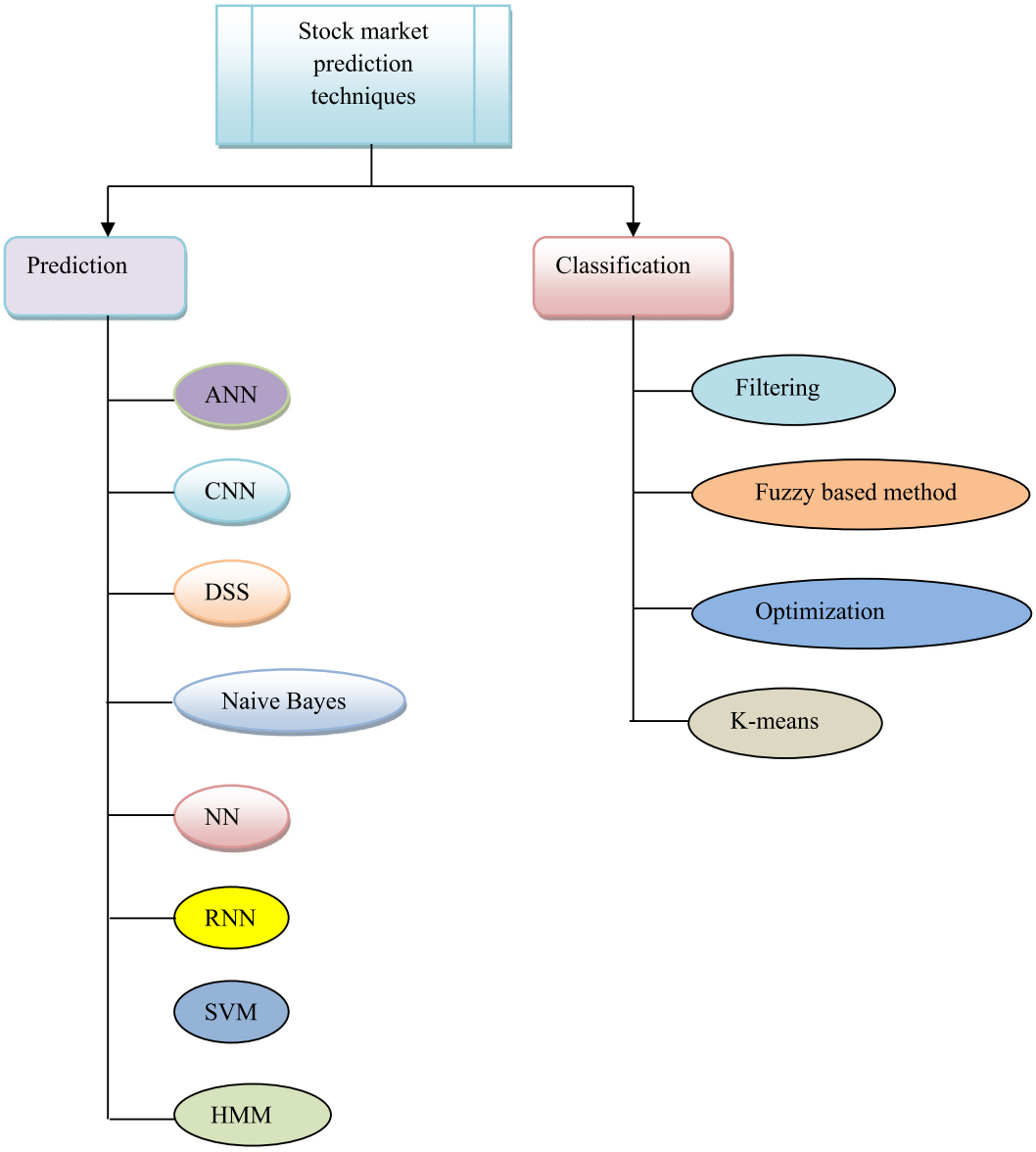


**Fig. 2.** Categorization of distinct stock market prediction techniques.

Mean Square (LMS) method along with the fitness function for training the networks using weights. The prediction quality was evaluated in terms of Mean Absolute Percentage Error (MAPE) using stock indexes. The inputs of the CEFLANN are selected from the previous stocks using varying market sectors and technical indicators for determining the optimal trends in stock prediction. Moreover, the optimal weights are selected by an adaptive Differential Evolution (DE) mechanism.

Shrivas, A.K. and Sharma, S.K [34] employed machine learning algorithms which involve SVM, CHAID, Classification, and Regression Technique (CART), and ANN to analyze and predict the stock trends using BSE SENSEX data. The method lacked advanced predictive models for predicting the stock market index, and this method was not applicable for other financial datasets, such as BSE 100, YAHOO. Chakravarty, S. and Dash, P.K [35] developed Functional Link Interval Type-2 Fuzzy Neural System (FLIT2FNS) to predict the stock market index. This model used a fuzzy rule base, which uses type-2 fuzzy sets in the antecedent part and results obtained from Functional Link Artificial Neural Network (FLANN) is in consequent part. Particle Swarm Optimization (PSO), Back propagation, and learning algorithms are utilized in an independent manner for optimizing the parameters of forecasting models. Zhong, X. and Enke, D [36] developed an effective mechanism based on data mining strategies to predict the daily trends of the S&P 500 Index return using the economic features. Three techniques, which include Fuzzy Robust Principal Component Analysis (FRPCA), Kernel-Based Principal Component Analysis (KPCA), and Principal Component Analysis (PCA) based on dimensionality reduction are used for the prediction. Advanced kernel functions along with relevant kernel parameters are required for the prediction.

Omidi, A. et al. [37] used NN using prices of stock returns from Iran Tractor Manufacturing Company based on the stock prices collected from two years. Here, NN, which is trained by employing back propagation algorithm for making an effective decision by classifying the input data. Hadavandi, E. et al. [38] developed an integrated method based on Genetic Fuzzy Systems (GFS) and ANN for constructing a proficient system based on stock price forecasting. Initially, a Stepwise Regression Analysis (SRA) was adapted for determining factors, which must influence the stock prices. In addition, the raw data was split into clusters using the Self-Organizing Map (SOM) NN. At last, the clusters

are subjected to GFS models using the rule base extraction and database tuning. Patel, J et al. [39] designed a method, named two-phase fusion approach, to predict the values of the future stock market. The first phase contains SVR, and the second phase contains Random Forest (RF), SVR, and ANN for the prediction. The outputs generated by the model are SVR–RF, SVR-SVR, and SVR–ANN and other prediction models. However, the method failed to incorporate a semi-supervised system for making it more robust and the prediction to be more accurate.

Patel, J. et al. [40] designed a layer, named Trend Deterministic Data Preparation Layer, in prediction models for analyzing the stocks. The method computes ten technical parameters utilizing the using stock trading data and represents the obtained technical indicators as a trend deterministic data. The method focused on short time prediction for estimating the values of stock returns. Moreover, the method lacks long term prediction for analyzing the quarterly stock prediction, profit returns, company's organizational stability, and revenues. Moghaddam, A.H et al. [41] designed ANN to predict the exchange rate of daily NASDAQ stocks. Several feed forward ANNs employed backpropagation algorithm for the training. The method takes previous stock prices for making the prediction. The method considers the daily stock of NASDAQ for testing the ability of model prediction. Guresen, E et al. [42] designed a model, named Generalized Autoregressive Conditional Heteroscedasticity (GARCH), for evaluating the NN for stock market prediction. The model evaluates hybrid neural networks, Dynamic Artificial Neural Network (DAN2) and Multi-Layer Perceptron (MLP), for extracting the input variables. Moreover, the method failed to discover if the GARCH, E-GARCH provide correct effects on forecasts in terms of correlated variables.

**(b) CNN based prediction techniques**

CNN is a feed-forward neural network. The number of hidden layers in a CNN is more than that in a conventional neural network. CNN is the renowned deep learning algorithms utilized to predict stock markets. This subsection discusses the research works adopting the CNN based stock market prediction as follows: Vargas, M.R et al. [43] used deep learning mechanism to predict the directional movement using Standard & Poor's 500 index with technical indicators. The deep learning methods are utilized for determining and analyzing complicated patterns in the data and allow to speed up the trading process. The method failed to adapt reinforcement learning algorithms for training the model on market simulations. Zhou, X et al. [44] developed a generic framework by adapting the LSTM and CNN for providing training to predict frequent stock market trends. The technique imitates the trading mode of the trader and uses training and testing set for analyzing the effects of the updated model cycle in predicting the performance. Moreover, the method lacks other predictive models under multiscale conditions. Xu, B et al. [45] developed a recurrent convolutional neural network for predicting the stock market trend. The network captured essential information from the stock market returns using the significant feature. In the first layer, an entity embedding layer was used to automate the learning process. The method failed to consider financial knowledge for optimizing the model for stock trend prediction.

**(c) DSS based prediction technique**

DSS can predict changes in stock prices, which are required by investors in the stock market. In this subsection, the stock market prediction technique based on DSS is discussed as follows,

Wen, Q et al. [46] developed an advanced intelligent trading system using oscillation box prediction by integrating the SVM algorithm with stock box theory. The box theory entails that a stock purchasing or retailing is successful if the price disrupts the original threshold value compared to other boxes. The trading mechanism using two bound forecasts is built for making effective decisions. The method failed to build advanced robust estimators for improving the accuracy of the forecasts by adopting additional soft computing methods.

**(d) HMM-based prediction techniques**

Recently, HMM is applied to forecast and predict the stock market. HMM is successful in analyzing and predicting time depending phenomena, or time series. The different researches adopting the HMM based stock prediction are elucidated in this subsection. Badge, J., [47] developed various macro-economic factors for Indian stock market with different macro-economic factors, like technical indicators. These technical indicators are employed for deciding the patterns of the market in a specific time. There exist multiple technical indicators for yielding effective forecast by applying Principal Component Analysis (PCA). The inputs taken for the stock market prediction are selected technical indicators. HMM is adapted for determining the future prices and is considered as a dominant stochastic model. Gupta, A., and Dhingra, B. [48] developed a Posteriori HMM approach for predicting the stock ethics using the previous data. This method considers fractional variations in stock indexes for training the HMM. Then, HMM is utilized for making the maximum Posteriori decision from the stock values. The method failed to consider correlations for constructing the model. The performance is improved using the quantization by taking hour-by-hour and minute-by-minute stock values.

**(e) Naive Bayes based prediction technique**

Naïve Bayes algorithm is a classification method, which generates Bayesian Networks for a given dataset based on Bayes theorem. It presumes that the given dataset contains a particular feature in a class, which is unrelated to any other feature. Naïve Bayes algorithm is easy to build and useful for very large datasets and outperforms highly sophisticated classification techniques. The research work employed with the Naive Bayes for the stock market prediction is deliberated as follows. More, A.M et al. [49] used Neuro-linguistic programming (NLP) approach for obtaining the stock information, and determining the stock charts, which can help users to determine the right investments with high profit. The Hadoop framework was used for accessing a huge amount of stock data parallel, and the Naive Bayes algorithm was used for making the decisions to deal with probability interference and the method used prior knowledge for predicting the future trends.

**(f) NN based prediction techniques**

NN is a series of algorithms that recognize the original relationships in a set of data through a process that mimics the way the human brain operates. This section elaborates the research works employing the NN based stock market prediction as follows: Chang, P.C [50] developed a model, named Evolving Partially Connected Neural Network (EPCNN), to forecast the stock trends with input as technical indicators. The architecture provides advanced features using different ANN features, in which the connection between the neurons is randomly, and there can be multiple layers. An evolutionary algorithm is adapted for improving the learning algorithm and training weights. EPCNN uses random connections between neurons for learning previous time series data for improving the neural networks. The method failed to incorporate advanced soft computing techniques for dealing with other time series data. Pang, X et al. [51] developed an innovative NN method to attain improved stock market predictions. The data is taken from the livestock market in real-time to analyze the stocks. Here, the deep LSTM based NN was designed using the embedded layer and the LSTM neural network using automatic encoder for predicting the stock trends. The embedded

layer and automatic encoder were used for vectorizing the data using LSTM. However, the method was not applicable to the European and American stock market with large accuracy.

Atsalakis, G.S et al. [52] designed a Wave Analysis Stock Prediction (WASP) system using the neuro fuzzy architecture that uses certain aspects of the Elliott Wave Theory. This theory was useful for attaining accurate stock market prediction. The method was complicated as the oscillator adapted was a slow-moving oscillator that was devised using the moving averages, which make the scenarios difficult. Chatzis, S.P. [53] designed solid forecasting mechanism for predicting the stock market under varying time frames. The method integrates various machine learning mechanisms using daily stocks from a huge spectrum of economies. The method applied a suite of machine learning method to select the appropriate variables from large sets. Lastly, a bootstrap sampling was adapted to adjust the imbalanced nature of the fitting dataset. The method failed to explore deep neural networks for setting high-frequency data with high accuracy. Shen, W. et al. [54] developed a Radial Basis Function Neural Network (RBFNN) for training the data and predict the stock indices using the Shanghai Stock Exchange. Artificial Fish Swarm Algorithm (AFSA) was employed to optimize the Radial Basis Function (RBF) in the learning process of RBF and was optimized by GA and PSO. The method failed to consider non-quantitative factors into mathematical algorithms for increasing the accuracy of stock market prediction.

Asadi, S et al. [55] developed a hybrid intelligent model for predicting the stock exchanges. The hybrid model is the integration of GAs, data reprocessing mechanisms, and Levenberg–Marquardt (LM) algorithm to learn the NN. The initial weight of NN is used for tuning the LM algorithm by adapting GA. The data pre-processing mechanisms contain input variable selection and data transformation for improving the overall model accuracy. The method is tested by applying the stock exchange indices for predicting the stock trends. Adebiyi, A.A. et al. [56] designed an upgraded predictive model for predicting the stock index using fuzzy neural network architectures. The effects of utilizing hybrid technical, fundamental, and market indicators are analyzed for initiating the prediction. The input variables are refined from market indicators and are subjected to the fuzzy-neural network to enhance the accuracy for predicting the stocks. However, the method failed to examine the critical impacts of expert opinion for predicting the quality of stock prices. Tsai, C.F et al. [57] designed classifier ensemble mechanism for analyzing the stock trends by examining the prediction performance. The ensembles are classified as follows, namely heterogeneous and homogeneous are designed. The ensembles are analyzed using average prediction accuracy. Finally, the homogeneous classifier ensembles use NNs for predicting the stock returns. However, the method failed to consider non-economic factors and other datasets for improving prediction accuracy.

**(g) RNN based prediction techniques**

An RNN is a class of ANN, in which connections between nodes form a directed graph along a temporal sequence. This allows it to exhibit temporal dynamic behavior. This subsection deals with different research works performing the stock market prediction by employing the RNN are as follows: Hsieh, T.J et al. [58] designed a united system, in which RNN based Artificial Bee Colony (ABC-RNN) algorithm are integrated for forecasting the stock prices. The model contains three phases, in which, initially, Haar wavelet is employed for decomposing the stock price time series data and to remove artifacts and noises. Secondly, the RNN was used for constructing the input features using Stepwise Regression–Correlation Selection (SRCS). Thirdly, ABC was adapted for optimizing the RNN weights and biases while designing the parameters. However, the method lacks advanced pattern selection mechanism for retrieving essential patterns from the data. Xie, X.K., and Wang, H [59] designed RNN for analyzing the time series data. The dataset consists of the huge amount of intraday data from China Shanghai Shenzhen 300 Index. The RNN was used to classify the daily features using intraday data. The performance of the model was analyzed using precision and average profit. Chen, W et al. [60] designed a model on the basis of RNN with Gated Recurrent Units (GRUs) for predicting the volatility of stocks in the Chinese stock market. The multiple price related features are subjected as an input to the model. However, the method failed to consider advanced machine learning method, like Interdependent Latent Dirichlet Allocation (ILDA), for making an accurate prediction.

**(h) SVM based prediction techniques**

SVM is the supervised learning model with associated learning algorithms that analyze data used for classification and regression analysis. SVM is a powerful predictive tool for stock predictions in the financial market. This subsection demonstrates the stock prediction techniques based on SVM classifiers.

Oztekin, A [61] designed a generic method for predicting the price of daily stock movements by combining three prediction models, which include adaptive neuro-fuzzy inference systems, ANN, and SVM, based on data analytics. The model was verified with ten-fold stratified cross-validation for minimizing the random sampling bias with improved accuracy. But the method failed to consider other stock markets for verifying the method in different countries. Porshnev, A et al. [62] designed an approach, named lexicon-based approach, for improving the accuracy of stock market indicators using the mindsets of Twitter users. The analysis of mindsets reveals the existence of eight crucial emotions in more than 755 million tweets. Moreover, SVM based techniques are applied for predicting DJIA and S&P500 indicators. However, the method failed to maximize the training periods and enhance sentiment analysis algorithms. Ni, L.P et al. [63] designed a model that combined SVM and fractal feature selection method for predicting stock price index. The fractal selection method was appropriate to solve the nonlinear issues and can spot the important features for the selection. The method failed to consider other factors, which are responsible for testing the prediction model that includes micro and macro factors in predicting the stocks. Zhang, X et al. [64] designed a multi-source multiple instance model that can integrate sentiments, events, and quantifiable data into an inclusive framework. The extraction and representation methods are used for capturing the news events. The news events and quantitative data influence stock fluctuations while predicting stocks market trends.

**(i) SVR based prediction techniques**

The SVR [65] utilizes the same principles as the SVM for classification, with only a few minor differences. This subsection demonstrates the techniques employing the SVR based stock prediction system.

Kazem, A et al. [66] designed a forecasting model on the basis of the firefly algorithm, SVR, and chaotic mapping for predicting prices of the stock market. The prediction model consists of three phases. The first phase deals with delay coordinate embedding method for reconstructing space dynamics. In the second phase, the chaotic firefly algorithm was adapted to optimize the SVR parameters. Lastly, the optimized SVR is adapted to predict the prices of stock markets. Moreover, the method used Structural Risk Minimization (SRM) in training the SVR process, for making the method more robust than the existing models. Kao, L.J et al. [67] designed a stock prediction model, which used Non-Linear Independent Component Analysis (NLICA) as pre-processing for extracting the features from variables. The features, named Independent Components (ICs), are subjected as

the input to SVR to build the prediction models. The method can determine the concealed information from the original data using the feature extraction process. Thus, NLICA helps to provide valuable information for financial processing. Xiong, T et al. [68] developed an algorithm, named Firefly Algorithm based Multi-Output Support Vector Regression (FA-MSVR) to determine the parameters of MSVR. The evaluation is done using economic criteria, cost of computation, and statistical criteria. Based on economic criteria, the performance is evaluated based on simple trading mechanism, by the relative forecast. Based on statistical criteria, the forecasting is performed using the forecast measures and testing methods. However, the method failed to consider other valued time series data to forecast the tasks, which include exchange rate.

**(2) Analysis based on clustering techniques**

In this section, the analysis is done based on clustering techniques developed in the literature and is described as follows,

**(a) Filtering based prediction techniques**

The filtering has attracted much attention and has a variety of possible applications in prediction. The distinct research works practicing the filtering based stock market prediction are deliberated in this section. Arévalo, R et al. [69] designed trading rules using flag pattern recognition and incorporates essential innovations using past researches. Initially, a dynamic window scheme is developed, which permits the stop loss and makes updated profit periodically. Moreover, the flag pattern follows the current trends and thus, an indicator named EMA was added for filtering the trades. The technical indicators are computed using 15-min and 1-day timeframes that facilitate simultaneous consideration of short and medium terms. But, the model failed to address data snooping problems after producing certain trading rules. Ariyo, A.A et al. [70] designed a predictive model using the ARIMA model for stock market prediction. The stock data published is obtained from the Nigeria Stock Exchange (NSE) and the New York Stock Exchange (NYSE). The results of the ARIMA models are used with the advanced forecasting techniques for short time prediction. Srinivasan, P. and Ibrahim, P [71] designed a forecasting model, using a GARCH model. This model was adapted to forecast the volatility of the SENSEX index obtained from a specific period. Using performance metrics, the GARCH model shows superior results in predicting the SENSEX Index return.

**(b) Fuzzy based prediction techniques**

Fuzzy logic is a form of many-valued logic, in which the truth values of variables may be any real number between 0 and 1. In this section, different researches adopting the fuzzy-based mechanisms for predicting the stock market is elucidated. Esfahanipour, A. and Aghamiri, W [72] designed a Neuro-Fuzzy Inference System for predicting the stock prices. The fuzzy model applied technical indexes as input, and the resulting part is a linear combination of the inputs. Fuzzy C-Means (FCM) clustering was adapted to identify the set of rules. The membership function was described using a Gaussian function. The purpose of ANFIS is to predict the variation between stock prices using several aspects, such as political reasons, macro-economic change, technical index, and fundamental analysis. Enke, D et al. [73] designed a three-phase stock market prediction system, in which the first stage covers multiple regression analysis for defining the financial variables. In the second phase, the differential evolution-based type-2 fuzzy clustering is adapted for creating a prediction model. In the third phase, a Fuzzy type-2 Neural Network is adapted for performing the reasoning for predicting future stock prices. Boyacioglu, M.A. and Avci, D [3] developed a model, named Adaptive Network-Based Fuzzy Inference System (ANFIS), for predicting the stock market returns. The purpose of the work is to determine if the ANFIS algorithm is able to predict the trends of the stock market return accurately.

Javedani Sadaei, H. and Lee, M.H [74] developed a multilayer model for predicting the stocks, including five layers. Each layer is taken into consideration for assisting forecast development by merging the problems completely. The method failed to consider the behavior of other layers discretely, which can be helpful for developing more layers for the systems. Chen, M.Y. and Chen, B.T [5] designed an advanced fuzzy time series model for predicting the prices of the stock market. The model was designed on the basis of granular computing approach using entropy-based discretization and binning-based partition mechanisms. The model adapted fuzzy time series models for predicting the stock indices. The method failed to consider other essential factor, which includes financial reports, technical indicators, and trading volume. Wei, L.Y et al. [11] developed a system named ANFIS, which used multiple technical indicators for predicting the trends of the stock price. At first, the required technical indicators are selected using advanced indicators with respect to the correlation matrix. Secondly, subtractive clustering mechanism is adapted, for partitioning the technical indicators using data discretization. Thirdly, Fuzzy Inference System (FIS) is used for extracting the rules of linguistic terms from technical indicators and optimizing the FIS parameters using adaptive networks for accurate prediction. However, the method was not applicable for predicting other stock indices. Other data discretization mechanism is required for evaluating the variation of performances. Sadaei, H.J et al. [75] designed an advanced fuzzy set for the fuzzy time series model and was based on the trend estimator for estimating the trends of data in an appropriate manner. Initially, the past data are fuzzified using different fuzzy sets for computing differential fuzzy relations. Lastly, the defuzzification is carried out to forecast the current stock market. The imperialist competitive algorithm was deployed for training the model to identify future trends with improved accuracy. The method failed to consider high-order differencing for dealing with seasonality.

**(c) K-means based clustering technique**

K-means clustering is a technique of vector quantization, which is fashionable for cluster analysis in data mining. The research paper utilizing the K-means based stock market prediction system is explained in this subsection: Nanda, S.R. et al. [76] designed a data mining mechanism for classifying the stocks into clusters. Once the classification is completed, the stocks were chosen from the groups for constructing a portfolio. The criterion for reducing the risk is attained by the portfolio diversification. The clustering mechanism divides stocks using specific investment criteria. The stock returns are used along with valuation ratios from BSE dataset. The clustering method was introduced for portfolio management and for choosing the stocks to attain effective frontier.

**(d) Optimization based clustering techniques**

Optimization is an act of deciding as fully perfect, functional, or effective as possible. The researches utilizing the optimization-based stock market prediction system are explained in this subsection. Cheng, C.H. et al. [77] designed a hybrid forecasting model that used multiple technical indicators for forecasting the trends of stock prices. The model employed four algorithms in the hybrid model for providing effective rules to forecast and use extracted rules along with support values using rough set theory. The first was the technical indicators, which are used to predict the upcoming stock prices based on correlation matrices. The second method employed was Cumulative Probability Distribution Approach (CDPA) for partitioning the technical indicators and third was Rough Set Theory (RST) algorithm for extracting the linguistic rules and fourth was GA for refining

the extracted rules to provide better prediction accuracy. But, the method failed to consider other data discretization methods and artificial intelligence algorithms for yielding improved forecasting process. AraúJo, R.D.A., and Ferreira, T.A [4] developed an Evolutionary Morphological-Rank-Linear (EMRLF) mechanism for overcoming the random walk to forecast financial time series. EMRLF contains intelligent hybrid model that integrates Modified Genetic Algorithm (MGA) and Morphological Rank Linear (MRL) filter to perform an evolutionary search for minimizing the total time lags that are capable for tuning the time series data in MRL filter. Then, the Least Mean Square (LMS) algorithm was employed for predicting the stock trends. The model was unable to regulate the distortions based on time while modeling financial time series data. Yeh, C.Y et al. [2] designed a two-phase multiple-kernel learning algorithm by adapting the gradient projection method and sequential minimal optimization. Based on the method, the benefits of various hyperparameter sets are devised to improve the overall performance of the system. In addition, the trial-and-error was used to determine suitable hyperparameter settings.

## 3. Research gaps and issues

In spite of the several distinct stock market prediction methodologies, there exist certain limitations, which must be addressed for attaining effective stock market prediction. This section deliberates the research gaps and issues in the different stock market prediction methods.

The challenges faced by the NN based on stock market prediction are as follows: The devised ANN was not declared as an effective scheme for predicting the stock market as the neural models cannot tolerate high computational overhead due to large neurons contained in the hidden layer and appropriate weight adaption [33]. NN, developed in [35], performed both the testing and the training at a slower rate; this affected the prediction performance. Moreover, overfitting, trapped in local minima and black box technique are the drawbacks which can be handled using NN. The obtained results of NN based stock market prediction system devised in [51] were with low accuracy due to the influence of the misclassification of the analogous patterns and the network parameters utilized were not optimized. The research issue in the CNN based stock prediction method is that the devised CNN with the deep learning framework was not suitable for the highly extensive applications. The rate of recognition accuracy achieved by CNN was comparatively poorer than the other state of the art prediction system for the stock prediction [45]. The devised decision support system [46] did not use the practical knowledge and techniques collected, for designing a workable stock expert system, in stock investment. ANN required prolonged training process for designing an optimal model and suffered from a lack of explanation for determining the solution is generated [61]. The fuzzy time series model [5] required more time and the accuracy computed by the fuzzy time series model for the prediction is perpetually affected due to length, and the formulation of splitting methods is very complicated, leading to complex computational process. In SVM, a feature selection method was adapted for predicting the trend of stock markets. The feature selection method was unable to point out the required number of optimal features, and thus, the accuracy of the system is highly affected. The devised SVM and NN based prediction system depend on the correlation value of the chosen feature [63].

The table contains 50 research papers, analyzed to forecast the trends in predicting the stock market. The techniques employed in the papers are categorized based on prediction and clustering techniques. The prediction technique is further classified based on ANN, CNN, DSS, HMM, NN, RNN, SVM, SVR, and Naive Bayes, and the clustering techniques are further classified based on filtering, optimization, fuzzy based techniques, and k-means techniques. Moreover, these papers are analyzed using certain factors, such as publication year, datasets employed, performance metrics, software tools, and the values of performance metrics. From the analysis, it is evaluated that more number of research papers were published in the year 2011. ANN was employed in most of the papers for predicting the market trends, whereas fuzzy based systems are employed for clustering the stock data. The most used software tool for predicting the market trends in stock prediction systems is MATLAB. The frequently used datasets for the stock market prediction are BSE, TAIEX, and S & P 500 indexes. Other datasets, such as NASDAQ, Istanbul stock exchange, Tehran stock exchange index, Citegroup and motors, Yahoo finance, Chinese stock market, National bank of Greece, and Nigeria stock exchange are also employed for the stock market prediction to analyze the market trends. The commonly used performance metrics are Mean Absolute Percent Error (MAPE), accuracy, Root-Mean-Square Error (RMSE), Mean Squared Error (MSE), Mean absolute error (MAE), and precision. Other metrics, such as sensitivity, specificity, Mean Percentage Error (MPE), Median Relative Absolute Error (MRAE), Average Relative Variance (ARV), Relative Difference in Percentage (RDP), Root Mean Squared (RMS), Coefficient of multiple determination (COV), F-measure, information gain, F1-Score, and Mathews Correlation Coefficient (MCC) are used in few works, for the analysis. The most used performance metric for analyzing the performance of stock market techniques is MAPE. Table 1 represents the analysis chart of the collected papers.

## 4. Analysis and discussion

This section explains the analysis of the stock market prediction based on publication year, adapted methods, datasets used, evaluation parameters, software tools, and accuracy.

### 4.1. Analysis on the basis of publication years

This subsection presents the analysis based on the publication years of the considered stock market prediction techniques. Fig. 3 illustrates the number of research papers published in the years from 2010 to 2018. From the 50 papers surveyed, more number of works i.e. 11 research papers were published in the year 2011. In 2018, seven research works were developed for stock market prediction.

### 4.2. Analysis based on prediction techniques

In this subsection, the analysis is carried based on the employed stock market prediction techniques. The techniques employed for effective stock market prediction is depicted in Fig. 4. From Fig. 4, it is noted that 29% of the works employed the ANN, 11% of the research papers used SVM, and 9% of the works are based on the SVR. The RNN is employed in 9% of the researches, decision support system is employed in 3% of the works, Naive Bayes is used in 3% and 23% of the works are based on NN, 6% papers are based on HMM and remaining 8% are based on CNN. Thus, NN and ANN are mostly employed techniques for stock market prediction.

### 4.3. Analysis based on clustering techniques

Fig. 5 presents the analysis on the basis of different clustering methods employed for the effective stock market prediction. Here, 56% of the research papers are based on Fuzzy based system, 19% of papers used Filtering, 19% of papers utilized optimization techniques, and 6% of papers adapted the K-means clustering. Thus, it can be noticed that the fuzzy-based technique was utilized in more research papers for stock market prediction.

**Table 1**
Analysis chart based on prediction and clustering techniques.

| Ref. no | Year | Performance evaluation measures | Datasets employed | Software tools | Prediction | Clustering |
|---|---|---|---|---|---|---|
| 1 | 2013 | MAPE | Microsoft Corp and Goldman Sachs | MATLAB | Bayesian based ANN | |
| 2 | 2014 | RMSE, MAPE | BSE, IBM Corp. | | ANN | |
| 3 | 2017 | Accuracy | DJIA | | | Filtering |
| 4 | 2016 | Accuracy, Sensitivity, and Specificity | BIST 100 Index | | SVM | |
| 5 | 2015 | MSE, RMSE, MAE, MPE | TAIEX | | | Fuzzy time series model |
| 6 | 2010 | Accuracy | TAIEX | | | Optimization |
| 7 | 2011 | RMSE | TAIEX | | | Fuzzy time series model |
| 8 | 2016 | MdRAE | NASDAQ, TAIEX | | | Fuzzy time series model |
| 9 | 2013 | MSE, MAPE ARV, POCID, NMSE | Petrobras Company Stock Prices, Directv Group Inc Stock Prices, and Microsoft Corporation Stock Prices, | | | Optimization |
| 10 | 2011 | RDP, RMSE | TAIEX | Multiple-kernel support vector regression | | Optimization |
| 11 | 2010 | RMS, R2, and Cov | ISE National 100 Index | | | Fuzzy system |
| 12 | 2018 | MAE and MAPE | BSE SENSEX dataset | | ANN | |
| 13 | 2010 | RMSE, MAPE | Tehran Stock Exchange Indexes | | | Fuzzy c means |
| 14 | 2012 | MAPE | Citigroup and Motors Liquidation Company | | Neural network | |
| 15 | 2012 | MAPE | S&P 500, BSE, and DJIA | | ANN | |
| 16 | 2018 | MSE, DA | Shanghai A-shares composite index | Python | Neural network | |
| 17 | 2010 | MSE , SCC | S&P500 | MATLAB | Decision support system | |
| 18 | 2010 | Intraclass inertia | Capitaline Databases Plus | MATLAB | | K-means |
| 19 | 2017 | SGD | S&P 500 index series | Tensorflow1. | CNN | |
| 20 | 2011 | Hit rate | National Bank of Greece index stock | | Neural network | Fuzzy system |
| 21 | 2016 | Accuracy | S&P 500 | MATLAB | ANN | |
| 22 | 2012 | MAPE | S&P CNX NIFTY | | HMM | |
| 23 | 2011 | RMSE, MAE, MAPE | DJIA, FTSE, Tokyo Nikkei-225 Index (Nikkei), and TAIEX | | RNN | |
| 24 | 2018 | G-mean, Accuracy | Full in-sample dataset. | | Neural Network | |
| 25 | 2010 | RMSE, MAPE | SENSEX Index returns of Indian stock market, | | | Filtering |
| 26 | 2011 | AER | Shanghai Stock Exchanges. | | Neural network | |
| 27 | 2011 | Minimum Squared Error | Iran Teraktorsazi Factory | MATLAB | ANN | |
| 28 | 2010 | MAPE | Dell Corporations from the IT sector and British airlines and Ryanair airlines from Airline sector., and daily the stock price of IBM. | | ANN | |
| 29 | 2012 | MAPE, POCID | TSE, TEPIX data | | Neural network | |
| 30 | 2011 | Accuracy | NSE | Matlab | Neural network | Fuzzy system |
| 31 | 2013 | Recall Precision F-measure | DJIA and S&P500 | Javascript | SVM | |
| 32 | 2013 | RMSE, MAD, MAPE, and DS | Nikkei 225 stock indexes, and Shanghai Stock Exchange Composite (SSEC) | | SVR | |

(*continued on next page*)

**Table 1** (*continued*).

| Ref. no | Year | Performance evaluation measures | Datasets employed | Software tools | Prediction | Clustering |
|---|---|---|---|---|---|---|
| 33 | 2011 | Accuracy | TEJ dataset | | Neural network | |
| 34 | 2014 | Accuracy | FTSE , S&P 500 for the US, 100 for the UK, and Nikkei 225 for Japan, | MATLAB | SVR | |
| 35 | 2014 | MAPE, MAE, rRMSE and MSE | S&P, BSE, CNX Nifty and Sensex from Indian stock markets | | ANN | |
| 36 | 2014 | Accuracy, Precision, Recall, F-measure | Reliance Industries and Infosys Ltd. and CNX Nifty | | ANN | |
| 37 | 2014 | RMSE | TAIEX, NASDAQ, DJIA, and S&P 500 | MATLAB | | Fuzzy time series |
| 38 | 2018 | RMSRE , DPA | China stock market | | CNN | |
| 39 | 2016 | Accuracy | NASDAQ stock exchange | MATLAB | ANN | |
| 40 | 2016 | Precision | China Shanghai Shenzhen 300 Index | | RNN | |
| 41 | 2012 | MAPE | Apple Inc., TATA steel, Dell Inc. and IBM Corporation | | HMM | |
| 42 | 2011 | RMSE | S&P 500 Index level | | | Fuzzy c means |
| 43 | 2011 | MSE and MAD | NASDAQ Stock Exchange index. | | ANN | |
| 44 | 2018 | Accuracy | | | Naïve Bayes algorithm | |
| 45 | 2017 | RMSEMAE, and MAPE | Chinese stock market | | RNN | |
| 46 | 2014 | Accuracy, Precision | NYSE and NSE | | | Filtering |
| 47 | 2011 | Symmetrical uncertainty, Information Gain | SSECI | | SVM | |
| 48 | 2012 | MSE MAPE | National Bank shares and Microsoft daily stock | MATLAB | SVR | |
| 49 | 2018 | Accuracy and MCC | Yahoo Finance | | CNN | |
| 50 | 2018 | F1 Score, Accuracy | Chinese dataset | | SVM | |

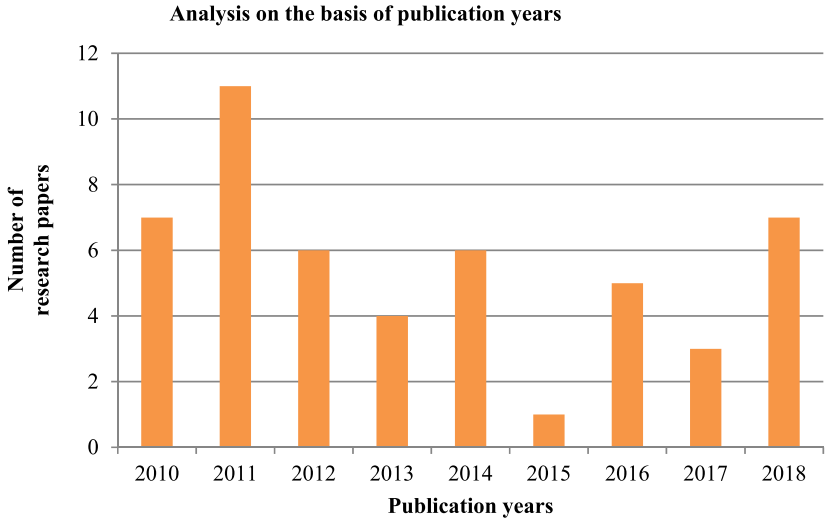


**Fig. 3.** Analysis on the basis of publication year.

### 4.4. Analysis based on a software tool

This section describes the analysis done on the basis of the software tool employed in the research papers. Table 2 deliberates the software tools adapted for performing the effective prediction of the stock market. The major software tools employed in the research papers are JavaScript, Python, MATLAB, and Tensor flow. From Table 2, it can be evaluated that the MATLAB is frequently used software tool for effective stock market prediction.

### 4.5. Analysis based on datasets

This section elaborates the analysis carried out based on the datasets adapted in the research works. Various datasets employed for the effective stock market prediction are depicted in

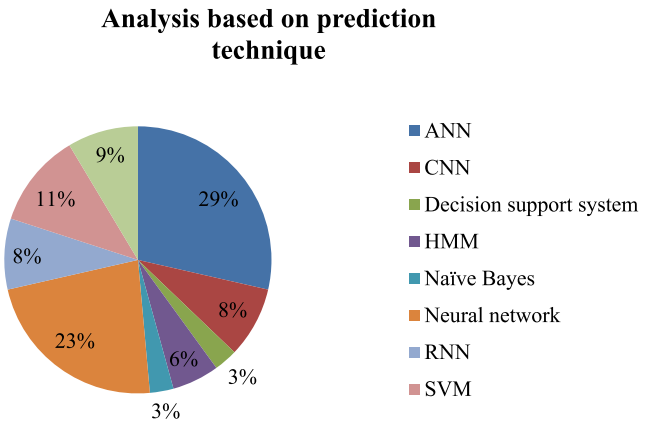


**Fig. 4.** Analysis based on prediction technique.

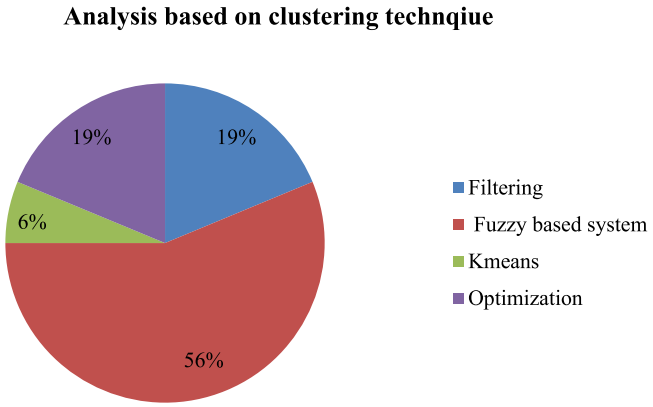


**Fig. 5.** Analysis based on clustering technique.

**Table 2**
Analysis based on the software tool.

| Implementation tools | Number of research papers |
|---|---|
| JavaScript | [62] |
| Python | [51] |
| MATLAB | [1,36,37,41,46,56,66,68,74,76] |
| Tensor flow | [43] |

Fig. 6. The frequently used datasets for the stock market prediction are BSE, TAIEX, and S & P 500 indexes. Other datasets considered are NASDAQ, Istanbul stock exchange, Tehran stock exchange index, Citegroup and motors, Yahoo finance, Chinese stock market, National bank of Greece, and Nigeria stock exchange.

### 4.6. Analysis based on performance metrics

The analysis carried out in terms of performance metrics for predicting the stock market is elaborated in this section using Table 3. The commonly used performance metrics are MAPE, RMSE, accuracy, MSE, MAE, and precision. Other metrics include sensitivity, specificity, MPE, MRAE, ARV, RDP,RMS,COV, F-measure, information gain, F1-Score, and MCC, which are employed in research papers [69], [69], [5], [75], [75], [2], [3], [3], [62], [63], [64], and [45], respectively.

### 4.7. Analysis based on values of performance metrics

The analysis on the basis of performance metrics value is described in this section. The analysis in terms of accuracy, MAPE, RMSE, and MSE is elaborated in the subsection.

**Table 3**
Analysis based on performance metrics.

| Performance metrics | Number of research papers |
|---|---|
| MAPE | [1,4,33–35,38,39,47,48,50,55,58,60,66,67,72] |
| RMSE | [2,5,11,33,39,58,60,67,72–75] |
| Accuracy | [36,40,41,45,49,51,53,56,57,59,64,68–70,77] |
| MSE | [4,5,39,42,46,47,51,66,72] |
| MAE | [5,34,39,58,60,72] |
| MRAE | [4,55] |
| NMSE | [4,72] |
| Hit rate | [52,56] |
| Recall | [40,62] |
| Precision | [40,59,62,70] |

**Table 4**
Analysis based on the accuracy.

| Accuracy range | Number of research papers |
|---|---|
| 50%–60% | [51,64] |
| 60%–70% | [11,45,57] |
| 70%–80% | [36,49,53,61,68] |
| 80%–90% | [40,56,70] |
| 90%–100% | [69] |

**Table 5**
Analysis based on MAPE.

| MAPE range | Number of research papers |
|---|---|
| $< 0.5$ | [4,33,35,47,56,58,66,72,73] |
| 0.5–1 | [34,50,60] |
| 1–1.5 | [1,38,71] |
| 1.5–2 | [40] |
| 2–2.5 | [67] |

**Table 6**
Analysis based on RMSE.

| RMSE range | Number of research papers |
|---|---|
| $< 0.3$ | [2,33,58,71,72,74,75] |
| 0.3–0.6 | [67,77] |
| 0.6–0.9 | [60] |
| 0.9–1.2 | [73] |

**(a) Analysis in terms of accuracy**

The analysis in terms of the accuracy values is elaborated in this section. Table 4 elaborates the analysis based on accuracy parameter specified using five ranges as 50%–60%, 60%–70%, 70%–80%, 80%–90% and 90% and 90%–100%. From the table, it is noted that the paper [69] attained improved accuracy with accuracy range 90%–100% and [40,56] and [70] had attained improved accuracy within the range 80%–90%, respectively. The minimum accuracy, i.e., accuracy range within 50% to 60%, is obtained by the research papers [51] and [64].

**(b) Analysis in terms of MAPE**

The analysis in terms of the MAPE values is described in this section. Table 5 elaborates the analysis based on MAPE parameter with five ranges as 0.0–0.5, 0.5–1.0, 1.0–1.5, 1.5–2, and 2–2.5, respectively. From the table, it is evaluated that the papers [1,40, 67,71] and [38] had attained improved MAPE within the range 2–2.5, 1.5–2 and 1–1.5, respectively.

**(c) Analysis in terms of RMSE**

The analysis in terms of RMSE values is described in this section. Table 6 elaborates the analysis in terms of RMSE parameter with four ranges as 0.0–0.3, 0.3–0.6, 0.6–0.9, and 0.9–1.2, respectively. From the table, it is elucidated that the papers [2,33, 58,71,72,75] and [74] had attained lower RMSE within the range $< 0.3$.

**(d) Analysis in terms of MSE**

The analysis in terms of the MSE values is described in this section. Table 7 elaborates the analysis in terms of MSE parameter

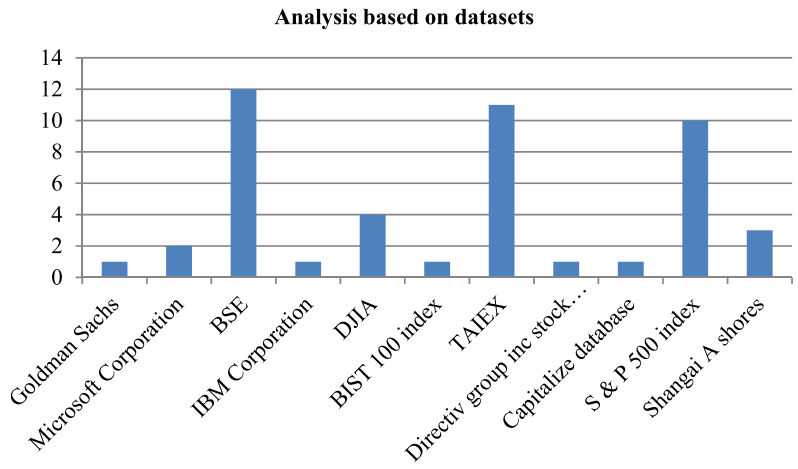


**Fig. 6.** Analysis based on datasets employed.

**Table 7**
Analysis based on MSE.

| MSE range | Number of research papers |
|---|---|
| < 0.007 | [2,35,45,46] |
| 0.007–1 | [5,37] |
| 1–2 | [39,42] |

with three ranges as < 0.007, 0.007–1, and 1–2, respectively. From the table, it is elucidated that the papers [2,35,46] and [45] had attained lower RMSE within the range < 0.007.

## 5. Conclusion

This paper presented a survey of different techniques employed for attaining effective stock market prediction, which is categorized on the basis of prediction techniques and clustering techniques. The purpose of this survey is to classify the existing techniques in terms of publication years, methodologies adapted, datasets used, performance metrics and implementation tools, using 50 research papers. The techniques utilized for the stock market prediction involves ANN, SVM, SVR, HMM, NN, fuzzy based techniques, K-means, and so on. In addition, the research gaps and the issues for predicting the stock market are elaborated for suggesting effective future scope. The commonly used technique for attaining effective stock market prediction is ANN and the fuzzy-based technique. These techniques can be effectively utilized for controlling and monitoring the entire stock market. The major challenge faced by the stock price prediction systems is that most of the existing techniques cannot be detected using historical stock data as they are affected due to certain factors, which involve government policy decisions, market sentiments and so on. Thus, the data from different sources are required for making decisions, and the data pre-processing is a complex task for data mining. These are major limitations that need to be addressed in the future by adapting advanced stock market prediction techniques. In the future, we will try to develop the method for providing the accurate prediction of the stock market.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

[1] J.L. Ticknor, A Bayesian regularized artificial neural network for stock market forecasting, Expert Syst. Appl. 40 (14) (2013) 5501–5506.
[2] C.Y. Yeh, C.W. Huang, S.J. Lee, A multiple-kernel support vector regression approach for stock market price forecasting, Expert Syst. Appl. 38 (3) (2011) 2177–2186.
[3] M.A. Boyacioglu, D. Avci, An adaptive network-based fuzzy inference system (ANFIS) for the prediction of stock market return: the case of the Istanbul stock exchange, Expert Syst. Appl. 37 (12) (2010) 7908–7912.
[4] R.D.A. AraúJo, T.A. Ferreira, A morphological-rank-linear evolutionary method for stock market prediction, Inform. Sci. 237 (2013) 3–17.
[5] M.Y. Chen, B.T. Chen, A hybrid fuzzy time series model based on granular computing for stock price forecasting, Inform. Sci. 294 (2015) 227–241.
[6] T. Kimoto, K. Asakawa, M. Yoda, M. Takeoka, Stock market prediction system with modular neural network, in: Proceedings of the International Joint Conference on Neural Networks, 1990, pp. 1–6.
[7] C. Nikolopoulos, P. Fellrath, A hybrid expert system for investment advising, Expert Syst. 11 (4) (1994) 245–250.
[8] K. Kim, I. Han, Genetic algorithms approach to feature discretization in artificial neural networks for prediction of stock index, Expert Syst. Appl. 19 (2000) 125–132.
[9] P. Vijaya, G. Raju, Santosh Kumar Ray, Artificial neural network-based merging score for Meta search engine, J. Comput. Sci. 12 (7) (2016) 323–340.
[10] T.H. Roh, Forecasting the volatility of stock price index, Expert Syst. Appl. 33 (2007) 916–922.
[11] L.Y. Wei, T.L. Chen, T.H. Ho, A hybrid model based on adaptive-network-based fuzzy inference system to forecast Taiwan stock market, Expert Syst. Appl. 38 (11) (2011) 13625–13631.
[12] M.R. Hassan, A combination of hidden Markov model and fuzzy model for stock market forecasting, Neurocomputing 72 (16–18) (2009) 3439–3446.
[13] R. Majhi, G. Panda, G. Sahoo, A. Panda, A. Choubey, Prediction of S & P 500 and DJIA stock indices using particle swarm optimization technique, in: Proceedings of the Proceeding of the IEEE Congress on Evolutionary Computation, 2008, pp. 1276–1282.
[14] Y. Kara, M. AcarBoyacioglu, O.K. Baykan, Predicting direction of stock price index movement using artifcial neural networks and support vector machines: The sample of the Istanbul Stock Exchange, Expert Syst. Appl. 38 (5) (2011) 5311–5319.
[15] M. Ghiassi, J. Skinner, D. Zimbra, Twitter brand sentiment analysis: a hybrid system using n-gram analysis and dynamic artifcial neural network, Expert Syst. Appl. 40 (16) (2013) 6266–6282.
[16] W. Huang, Y. Nakamori, S.-Y. Wang, Forecasting stock market movement direction with support vector machine, Comput. Oper. Res. 32 (10) (2005) 2513–2522.
[17] A.F.S. Elsir, H. Faris, A comparison between regression, artifcial neural networks and support vector machines for predicting stock market index, Int. J. Adv. Res. Artif. Intell. 4 (7) (2015).
[18] S.S. Appadoo, Pricing Financial Derivatives with Fuzzy Algebraic Models: A Teoretical and Computational Approach, University of Manitoba, Winnipeg, 2006.
[19] A. Tavaneswaran, K. Tiagarajah, S.S. Appadoo, Fuzzy coefcient volatility models with applications, Math. Comput. Modelling 45 (7–8) (2007) 777–786.
[20] C. Carlsson, R. Fuller, On possibilistic mean value and variance of fuzzy numbers, Fuzzy Sets and Systems 122 (2) (2001) 315–326.

[21] A. Tavaneswaran, S.S. Appadoo, A. Paseka, Weighted possibilistic moments of fuzzy numbers with applications to GARCH modeling and option pricing, Math. Comput. Modelling 49 (1–2) (2009) 352–368.

[22] X. Zhou, Z. Pan, G. Hu, S. Tang, C. Zhao, Stock market prediction on high-frequency data using generative adversarial nets, Math. Probl. Eng. (2018).

[23] Manna Majumder, M.D. Anwar Hussian, Forecasting of Indian Stock Market Index using Artificial Neural Network, National Stock Exchange of India Limited, 2009.

[24] M.M.P. Bharambe, S.C. Dharmadhikari, Stock market analysis based on artificial neural network with big data, in: Proceedings of 8th Post Graduate Conference for Information Technology, 2017.

[25] U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, The KDD process for extracting useful knowledge from volumes of data, Commun. ACM 39 (11) (1996) 27–34.

[26] A. Kumar, Artificial neural networks for data mining, IASRI, Library Avenue, Pusa, New Delhi Google Scholar.

[27] D. Enke, S. Thawornwong, The use of data mining and neural networks for forecasting stock market returns, Expert Syst. Appl. 29 (4) (2005) 927–940.

[28] U.M. Fayyad, D. Haussler, P.E. Stolorz, KDD for science data analysis: Issues and examples, in: KDD, 1996, pp. 50–56.

[29] PhichhangOu, Prediction of stock market index movement by ten data mining techniques, Mod. Appl. Sci. 3 (12) (2009).

[30] A. Sharma, D. Bhuriya, U. Singh, Survey of stock market prediction using machine learning approach, in: Proceedings of International Conference on Electronics, Communication and Aerospace Technology, Vol. 2, 2017, pp. 506–509.

[31] J.R. Woolridge, C.C. Snow, Stock market reaction to strategic investment decisions, Strateg. Manage. J. 11 (5) (1990) 353–363.

[32] S. Arun Joe Babulo, et al., Stock market indices prediction with various neural network models, Int. J. Comput. Sci. Mobile Appl. 2 (3) (2014) 42–46.

[33] A.K. Rout, B. Biswal, P.K. Dash, A hybrid FLANN and adaptive differential evolution model for forecasting of stock market indices, Int. J. Knowl.-Based Intell. Eng. Syst. 18 (1) (2014) 23–41.

[34] A.K. Shrivas, S.K. Sharma, A Robust Predictive Model for Stock Market Index Prediction using Data Mining Technique, 2018.

[35] S. Chakravarty, P.K. Dash, A PSO based integrated functional link net and interval type-2 fuzzy logic system for predicting stock market indices, Appl. Soft Comput. 12 (2) (2012) 931–941.

[36] X. Zhong, D. Enke, Forecasting daily stock market return using dimensionality reduction, Expert Syst. Appl. 67 (2017) 126–139.

[37] A. Omidi, E. Nourani, M. Jalili, Forecasting stock prices using financial data mining and Neural Network, in: Proceedings of 3rd International Conference on Computer Research and Development, Vol. 3, 2011, pp. 242–246.

[38] E. Hadavandi, H. Shavandi, A. Ghanbari, Integration of genetic fuzzy systems and artificial neural networks for stock price forecasting, Knowl.-Based Syst. 23 (8) (2010) 800–808.

[39] J. Patel, S. Shah, P. Thakkar, K. Kotecha, Predicting stock market index using fusion of machine learning techniques, Expert Syst. Appl. 42 (4) (2015) 2162–2172.

[40] J. Patel, S. Shah, P. Thakkar, K. Kotecha, Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques, Expert Syst. Appl. 42 (1) (2015) 259–268.

[41] A.H. Moghaddam, M.H. Moghaddam, M. Esfandyari, Stock market index prediction using artificial neural network, J. Econ. Financ. Adm. Sci. 21 (41) (2016) 89–93.

[42] E. Guresen, G. Kayakutlu, T.U. Daim, Using artificial neural network models in stock market index prediction, Expert Syst. Appl. 38 (8) (2011) 10389–10397.

[43] M.R. Vargas, B.S. de Lima, A.G. Evsukoff, Deep learning for stock market prediction from financial news articles, in: Proceedings of IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications, 2017, pp. 60–65.

[44] X. Zhou, Z. Pan, G. Hu, S. Tang, C. Zhao, Stock market prediction on high-frequency data using generative adversarial nets, Math. Probl. Eng. (2018).

[45] B. Xu, D. Zhang, S. Zhang, H. Li, H. Lin, Stock market trend prediction using recurrent convolutional neural networks, in: Proceedings of CCF International Conference on Natural Language Processing and Chinese Computing, 2018, pp. 166–177.

[46] Q. Wen, Z. Yang, Y. Song, P. Jia, Automatic stock decision support system based on box theory and SVM algorithm, Expert Syst. Appl. 37 (2) (2010) 1015–1022.

[47] J. Badge, Forecasting of indian stock market by effective macro-economic factors and stochastic model, J. Stat. Econom. Methods 1 (2) (2012) 39–51.

[48] A. Gupta, B. Dhingra, Stock market prediction using hidden markov models, in: Proceedings of Students Conference on Engineering and Systems, 2012, pp. 1–4.

[49] A.M. More, P.U. Rathod, R.H. Patil, D.R. Sarode, B.E. Student, Stock market prediction system using hadoop, Internat. J. Engrg. Sci. (2018) 16138.

[50] P.C. Chang, A novel model by evolving partially connected neural network for stock price trend forecasting, Expert Syst. Appl. 39 (1) (2012) 611–620.

[51] X. Pang, Y. Zhou, P. Wang, W. Lin, V. Chang, An innovative neural network approach for stock market prediction, J. Supercomput. (2018) 1–21.

[52] G.S. Atsalakis, E.M. Dimitrakakis, C.D. Zopounidis, Elliott Wave Theory and neuro-fuzzy systems, in stock market prediction: The WASP system, Expert Syst. Appl. 38 (8) (2011) 9196–9206.

[53] S.P. Chatzis, V. Siakoulis, A. Petropoulos, E. Stavroulakis, N. Vlachogiannakis, Forecasting stock market crisis events using deep and statistical machine learning techniques, Expert Syst. Appl. 112 (2018) 353–371.

[54] W. Shen, X. Guo, C. Wu, D. Wu, Forecasting stock indices using radial basis function neural networks optimized by artificial fish swarm algorithm, Knowl.-Based Syst. 24 (3) (2011) 378–385.

[55] S. Asadi, E. Hadavandi, F. Mehmanpazir, M.M. Nakhostin, Hybridization of evolutionary Levenberg–Marquardt neural networks and data pre-processing for stock market prediction, Knowl.-Based Syst. 35 (2012) 245–258.

[56] A.A. Adebiyi, C.K. Ayo, S.O. Otokiti, Fuzzy-neural model with hybrid market indicators for stock forecasting, Int. J. Electron. Finance 5 (3) (2011) 286–297.

[57] C.F. Tsai, Y.C. Lin, D.C. Yen, Y.M. Chen, Predicting stock returns by classifier ensembles, Appl. Soft Comput. 11 (2) (2011) 2452–2459.

[58] T.J. Hsieh, H.F. Hsiao, W.C. Yeh, Forecasting stock markets using wavelet transforms and recurrent neural networks: An integrated system based on artificial bee colony algorithm, Appl. Soft Comput. 11 (2) (2011) 2510–2525.

[59] X.K. Xie, H. Wang, Recurrent neural network for forecasting stock market trend, in: Proceedings of International Conference on Computer Science, Technology and Application, 2017, pp. 397–402.

[60] W. Chen, Y. Zhang, C.K. Yeo, C.T. Lau, B.S. Lee, Stock market prediction using neural network through news on online social networks, in: Proceedings of International Conference on Smart Cities, 2017, pp. 1–6.

[61] A. Oztekin, R. Kizilaslan, S. Freund, A. Iseri, A data analytic approach to forecasting daily stock returns in an emerging market, European J. Oper. Res. 253 (3) (2016) 697–710.

[62] A. Porshnev, I. Redkin, A. Shevchenko, Machine learning in prediction of stock market indicators based on historical data and data from twitter sentiment analysis, in: Proceedings of 13th International Conference on Data Mining Workshops, 2013, pp. 440–444.

[63] L.P. Ni, Z.W. Ni, Y.Z. Gao, Stock trend prediction based on fractal feature selection and support vector machine, Expert Syst. Appl. 38 (5) (2011) 5569–5576.

[64] X. Zhang, S. Qu, J. Huang, B. Fang, P. Yu, Stock market prediction via multi-source multiple instance learning, IEEE Access (2018).

[65] C.A. Ghuge, Sachin D. Ruikar, V. Chandra Prakash, Support vector regression and extended nearest neighbor for video object retrieval, Evol. Intell. (2018) 1–14.

[66] A. Kazem, E. Sharifi, F.K. Hussain, M. Saberi, O.K. Hussain, Support vector regression with chaos-based firefly algorithm for stock market price forecasting, Appl. Soft Comput. 13 (2) (2013) 947–958.

[67] L.J. Kao, C.C. Chiu, C.J. Lu, J.L. Yang, Integration of nonlinear independent component analysis and support vector regression for stock price forecasting, Neurocomputing 99 (2013) 534–542.

[68] T. Xiong, Y. Bao, Z. Hu, Multiple-output support vector regression with a firefly algorithm for interval-valued stock price index forecasting, Knowl.-Based Syst. 55 (2014) 87–100.

[69] R. Arévalo, J. García, F. Guijarro, A. Peris, A dynamic trading rule based on filtered flag pattern recognition for stock market price forecasting, Expert Syst. Appl. 81 (2017) 177–192.

[70] A.A. Ariyo, A.O. Adewumi, Ayo, Stock price prediction using the ARIMA model, in: Procedings of Computer Modelling and Simulation, 2014, pp. 106–112.

[71] P. Srinivasan, P. Ibrahim, Forecasting stock market volatility of BSE-30 index using GARCH models, Asia Pac. Bus. Rev. 6 (3) (2010) 47–60.

[72] A. Esfahanipour, W. Aghamiri, Adapted neuro-fuzzy inference system on indirect approach TSK fuzzy rule base for stock market analysis, Expert Syst. Appl. 37 (7) (2010) 4742–4748.

[73] D. Enke, M. Grauer, N. Mehdiyev, Stock market prediction with multiple regression, fuzzy type-2 clustering and neural networks, Procedia Comput. Sci. 6 (2011) 201–206.

[74] H. JavedaniSadaei, M.H. Lee, Multilayer stock forecasting model using fuzzy time series, Sci. World J. (2014).

[75] H.J. Sadaei, R. Enayatifar, M.H. Lee, M. Mahmud, A hybrid model based on differential fuzzy logic relationships and imperialist competitive algorithm for stock market forecasting, Appl. Soft Comput. 40 (2016) 132–149.

[76] S.R. Nanda, B. Mahanty, M.K. Tiwari, Clustering Indian stock market data for portfolio management, Expert Syst. Appl. 37 (12) (2010) 8793–8798.

[77] C.H. Cheng, T.L. Chen, L.Y. Wei, A hybrid model based on rough sets theory and genetic algorithms for stock price forecasting, Inform. Sci. 180 (9) (2010) 1610–1629.