Section: Research Paper

Bibliometric Reviews of Stock Market Prediction: A Comprehensive Approach

***Priyanka Mahajan,**
Department of Computer Engineering and Technology
Guru Nanak Dev University
Amritsar, Punjab, India
p_pankaj_gupta@yahoo.co.in
ORCID: 0000-0002-6040-8943
Prabhpreet Kaur
Department of Computer Engineering and Technology
Guru Nanak Dev University
Amritsar, Punjab, India
prabhpreet.cst@gndu.ac.in
ORCID: 0000-0001-8498-5940

DOI:10.48047/ecb/2023.12.si4.764

Abstract

Stock market forecasting has always been a significant task. Its importance lies in the fact that predicting stock rates successfully helps us gain attractive profits through wise decisions. Although the prediction helps us earn profits yet, it's a major challenge due to blaring, uncertain, non-stationary, and non-linear financial market behaviour. Many techniques help to predict stock market trends. This study reviewed approximately 100 research papers that suggested methodologies, like Artificial Neural Networks (ANN), Support Vector Machine (SVM), Fuzzy Classifiers, Machine Learning Methods, and so on, based on stock market prediction. Few efforts have been undertaken to list the research gaps and the challenges faced by the existing techniques, which help the researchers to upgrade the future works. The works are analyzed using the bibliometric approach carried out between 1975 and 2021 using Scopus data, thereby highlighting the citation patterns and involving journals, authors, and countries in forecasting the stock market. Furthermore, keywords analysis is also staged using VOSViewer and Bibliometrix tools to visualize the research patterns. Despite a lot of research efforts, the current stock market prediction techniques still have many limits. This survey finally, tries to conclude that the stock market prediction is a very complex task, and different factors should be taken into account for predicting the future of the market more accurately and efficiently.

Keywords *stock market prediction, Technical analysis, machine learning, bibliometrics and systematic review analysis,*

1 Introduction

In today's modern era, the financial market has emerged as a pivot field in every growing and developing economy, country or society, since it influences the economic development of the country worldwide ([1]**;**[2]**).** Stock market plays an important role in world economy because financial development of any country is directly proportional to its financial activities[3]. The financial markets comprise of stocks, derivatives, currencies, bonds, commodities, etc. which are explored, investigated, studied, and traded in a detailed manner[4]; and this trading includes buying and selling of all the above listed instruments. In India, the stock market is controlled by a regulatory body known as Securities and Exchange Board of India (SEBI).

Stock market is such market which enables continuous exchange of buying and selling of companies stocks. Stock market, a subcategory from financial market, also called as equity market is a place where the new issues

of stocks, i.e., initial public offerings (IPOs) are enlisted and sold in the primary market and the succeeding buying and selling of stocks are carried out in the secondary market [5]**.** A stock can be bought or sold only if it is listed on any exchange. A stock exchange makes easy for stock brokers to trade company's stocks and other securities. Examples of some of the leading exchanges around the world include -- NYSE, NASDAQ of US, and London Stock Exchange of United Kingdom etc. The increases in salaries of people have made financial investments gain significant importance. In the past time people used to bank their savings, but due to low deposit rates and high inflation rates this proved a poor strategy for investments [6]. Among various financial tools such as banks and funds, investors are turning towards stock market investments due to the higher profits and returns. However, higher returns are associated with higher risks as well [6].

The privilege provided by stock market is that it yields higher dividends and profits as compared to other financial investments; while the negative of this market is higher risks associated with it, although this risk can be lowered by smart decisions. The experts, analysts and researchers believe that stock market is gaining significance due to the exponential growth in interest of people trading in it [7]**.** The main motivation behind investments in a stock market is mainly to gain potential benefits from these investments [8]. This reason forced stock market forecasters focus on developing methods to successfully forecast/predict index values or stock prices.[9] Forecasting of stock market trends is designated as a significant task, as predicting stock prices successfully may lead to attractive higher risk adjusted returns; from proper decisions. Since stock market is highly chaotic, volatile, dynamic, intrinsic noisy and uncertain field hence, its forecasting has become a burning topic of financial world these days.

The uncertainty in market trends, that is, the rise and fall in stock prices from the stock markets with respect to changes in foreign investments has made the task of prediction extremely challenging. Hence, the prediction task to purchase the stocks which bring profits to a person and to sell those stocks which are probably to fall in price has become really tedious. Although careful trading may earn higher returns, the risk associated with it may sometimes result in the losses unpredictably. From many years, investors and researchers are involved in developing and testing models of stock price prediction because this market has a large impact on businesses, education, jobs and economy[3]. The central motive behind successful stock market prediction [SMP] is achieving best results using minimum required input data.

1.1 Terms used in Stock Market

a. Stock Market

A stock market is also known as equity market where shares trading take place [10]. Shares are units of equity ownership in any corporation. The capital market comprises of Primary market as well as Secondary market.

b. Stock Exchange

Stock exchange serves as the trading platform that facilitates regular buying and selling of the listed shares. The stock exchange acts as a facilitator for the capital-raising process and receives a fee for its services from the company and its financial partners. Shares are either directly supplied by company through Initial Public Offer (IPO) as in primary market or can be bought from stock exchange as in secondary market.

c. *Stock Index*

Every Stock Exchange has its own Stock Index value. A stock market index helps to measure a stock market. An index collects data from a variety of companies across industries. There are thousands of companies listed on stock markets, making it extremely difficult to monitor each company. This is why stock market indices are created. Market indices make a group of company stocks and regularly measure them to reflect the performance of the overall market or a certain segments of the market. In other words, an index helps investors understand the health of the stock market, enabling them to study the market sentiment. For example, Dow Jones Industrial Average (DJIA), the Standard & Poor's 500 (S&P 500) and the NASDAQ Composite are the three major stock indexes of US. Similarly, S&P BSE Sensex and Nifty 50 are two major stock indices of BSE (Bombay Stock Exchange) and NSE (National Stock Exchange) respectively from Indian stock market [10].

1.2 Unpredictability of stock market movement

Although efficient market hypothesis (EMH) established by Malkiel and Fama (1970)[11], later revised by Fama (1991) [12] follow stochastic patterns making the list of stock prices highly dynamic, complicated, noisy, nonparametric, and nonlinear by nature. The complexity of the stock market is also associated with ample of factors such as political events, market news, quarterly earnings reports, international influence and conflicting trading behavior [13], therefore making it highly unpredictable.

Hence if it had been easy for investors and analysts to predict the chaotic behaviour of market with accuracy, this would have enabled them to earn profitable returns from the market [14]**.** Although, the accurate forecasting of financial markets for analyzing the trends is quite complex due to intrinsic noisy environments, uncertainty, random fluctuations with respect to the market trends[15]**,** yet **s**everal methods have been proposed in economics as well as in computer science to predict futuristic behaviour of market. Some of these methods are predicting stock trend direction (rise or fall, i.e., bull market or bear market, respectively), predicting intraday or inter-day stock prices, risks and returns associated with them and many more [4]**.**

This concluded that any predictive model which has the capability of generating high profits after predicting the market indices over time; proves itself not only as a strong evidence in opposition to the theory of EMH, but would also gather large profits with financial operations. The search for prediction models and profitable systems is still attracting a lot of attention from researchers [16]. These prediction models deploy two different types of analytical approaches named as Fundamental analysis (FA) and technical analysis (TA) as shown in figure 1.

2. Stock Market Prediction Approaches

The classical financial market prediction techniques are (i) Fundamental analysis and (ii) Technical analysis, also called as chartist analysis. FA and TA are also known as important decision making tools for stock market prediction. As stated in figure 1 these approaches are easily differentiated on the basis of information set used for prediction and decision making. [17].

The first approach i.e. fundamental analysis considers fundamental information as significant tool on which the investor believes for accurate predictions of stock market. The investor looks on a company's turnover, balance-sheet, expenses, income statements, annual and quarterly reports, profit and loss, assets and liabilities etc. [14] price to earnings ratio (P/E ratio) as well as it also includes macroeconomic time series variables such as, GDP (Gross Domestic Product), CPI (Customer Price Index), Currency Exchange Rates[18] to make decisions for investments. The price-to-earnings ratio (P/E) is one of the most widely used metrics by investors and analysts to determine stock valuation. Four fundamental indicators namely price earnings ratio (P/E), dividend payout ratio return on equity (ROE) and book value were applied by Vanstone [19] for forecasting of Australian stock market.

Few researchers selected some goods stocks for investment after comparing the financial ratios of many stocks. Their results were based on the principle of fundamental analysis. These researchers included Dutta, Avijan, Gautam and Sengupta [100]. The authors analysed their one-year return and compared with the predicted value -- i.e., Nifty—which gives an accuracy of 74.6%. This research study is one of the arts, which considers applying fundamental features to identify stocks for investments. Long term investors generally prefer fundamental analysis to earn profits from stock market [10]

The randomness of stock markets is clearly visible when it is looked at from outside of the field; however the researchers and investors predict the movements of the market and future price of stocks by analyzing previous trends and patterns of stock price levels and volume on candlestick charts. This method of prediction after analyzing the stock price and volume data from the previous history, to predict the future movements of the stock price is termed technical analysis [20]. The technical analysis uses stock charts and some set of tools called technical Indicators (TI) for identifying patterns and trends for evaluation of those stocks which can lead for profits. Various research studies ([21]; [22]; [23])) and many more like ([24];[25];[26]; [27]; [28]) proved that future stock-price movement can be predicted through use of technical analysis.

The uncertain behavior of stock market is also influenced by various macro-economical factors which include economic condition of any country or the company, bank rate, currency exchange rate, gold price, commodity price, movement of other stock market, psychology of investors, expectations of investors, political events, company's policies etc. ([1],[29]). Boyacioglu et al. [30] used various macro-economic factors such as "industrial production index", "consumer price index", "interest rates on deposits", "US Dollar exchange rates", "republic gold selling price", "interest rates" "closing price" of 4 indices namely DJIA, ISE National 100.

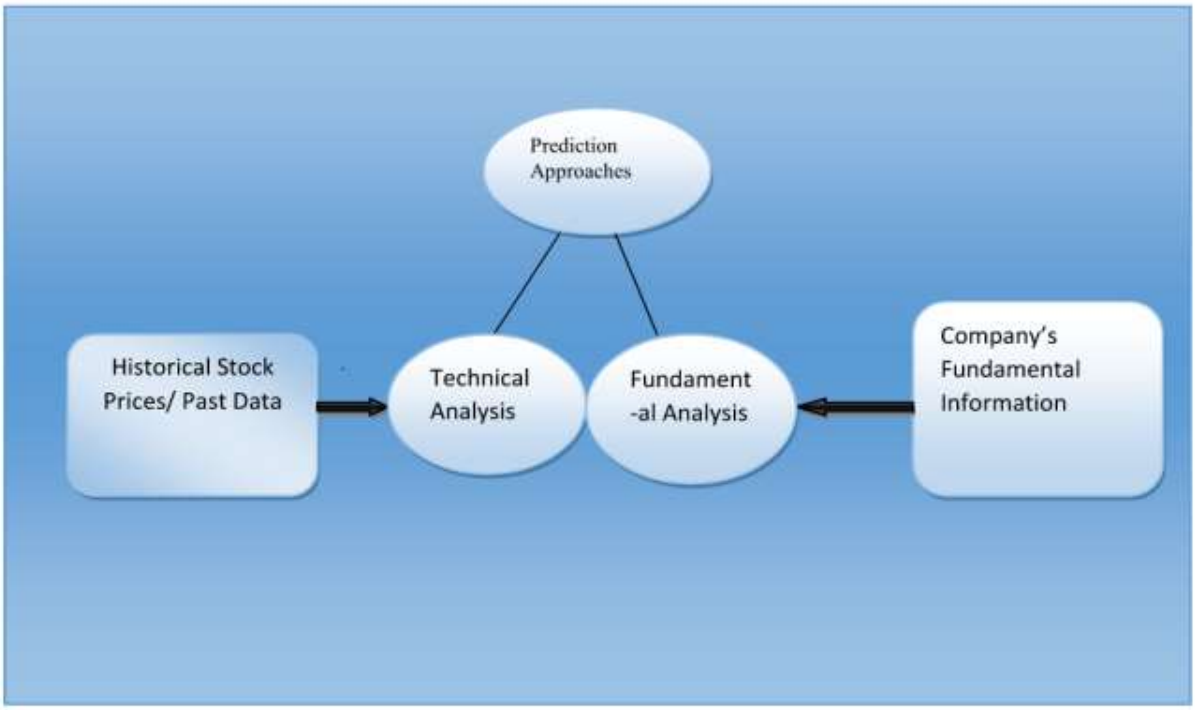

Figure1 Approaches for Stock market prediction

The prediction of stock prices and market indices also use a third approach, i.e. time series analysis which is /further divided into linear and nonlinear models. These statistical or Linear models such as the Autoregressive Integrated Moving Average (ARIMA) and the Autoregressive Conditional Heteroskedasticity (ARCH), and generalized autoregressive conditional heteroskedasticity (GARCH) [31], Kalman filtering, discriminating analysis etc. have been used to predict the stock market in finance, and have made some achievements. However, they had a limitation that they assumed a linear and stationary time series, which is inconsistent with the dynamic, non-linear characteristics of the real stock market. Hence they proved inefficient for stock market prediction.

In a review performed by Kumar and Ravi [32]**,** 128 papers about bankruptcy prediction of banks and firms were studied. This review shows that ANN methods outperform many methods and hybrid systems can combine the advantages of different methods. Primary studies from 2009 to 2015 examined by Cavalcante et al. [33], based on computational intelligence techniques for forecasting prices in the stock market, it was identified that TIs play a crucial role in SMP (Stock market Prediction). Although it becomes an overhead while making decision of choosing an appropriate TI in forecasting as the list of TIs is lengthy as well as variable [17]**.** Later, with the introduction of artificial intelligence (AI) and soft computing, these techniques have received increased attention within stock market prediction studies. Unlike traditional time series methods, these techniques can handle the nonlinear, chaotic, noisy, and complex data of the stock market, leading to more effective predictions [34]**.** Due to the advancement of computational power in recent times, predicting the stock market has been much faster and accurate. Artificial Intelligence and machine learning models play a crucial role in predicting the stock prices and, hence, determining an accurate result [35]**.**

Since stock data can be categorized as non-stationary, non-linear time series based data; machine learning techniques have been used for its prediction. Artificial Neural Networks (ANN) and Support Vector Machine

(SVM) are two machine learning algorithms which are most widely used for predicting stock and stock price index movement. Figure 2 illustrates deeply various categories of stock market prediction techniques.

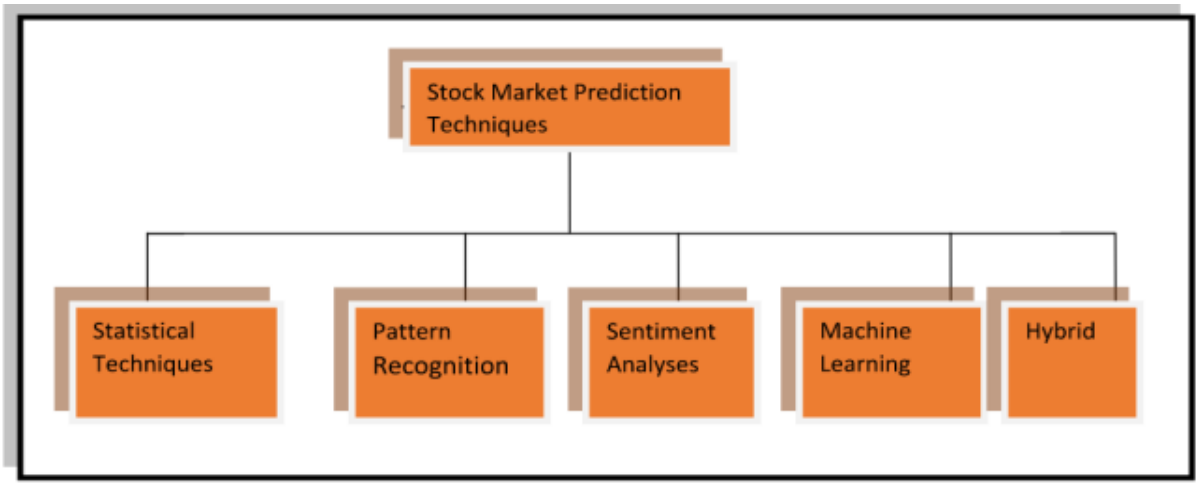

Figure Various Stock Market Prediction Techniques

Nowadays, stock market movements are analysed and predicted by the above mentioned four categories—statistical, pattern recognition, and machine learning (ML), and sentiment analysis.

2.1 Statistical Techniques

Statistical techniques which assume linearity, stationarity, and normality provided a way to analyse and predict stocks long time before machine learning techniques paved the way for prediction of stock price movements. Time series analysis is a statistical method that analyses and manipulates time series data. Time series is made of data points collected at constant time intervals. Zhong and Enke[36] defined, one group of statistical approaches, which used time series as input variables; the algorithms under this category included the Auto-Regressive Moving Average (ARMA), the Auto-Regressive Integrated Moving Average (ARIMA) and the Smooth Transition Autoregressive (STAR) model.

The ARIMA model is a widely used technique for stock market analysis [37]**.** They also described second group of statistical approaches which relies on utilizing multiple input variables, some of which are Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA), and regression algorithms.

2.2 Pattern Recognition

This is the other method for SMP; which is named as pattern recognition techniques. It showed own results but results obtained on stock prediction were not satisfactory [38]. These techniques gave excellent results while analysing and mining patterns rather than predicting the actual values.

2.3 Machine Learning Approach

The philosophy behind machine learning is to extract knowledge from data [39]**.** This is why machine learning has become extremely popular for the prediction of financial markets. Machine learning tasks are of two categories supervised and unsupervised learning. In supervised learning –"A set of labelled input data for training the algorithm and observed output data are available". On the contrary, the unsupervised learning contains only the unlabelled or observed output data. The supervised learning methods train an algorithm to automatically map the input data to the given output data. When trained, the machine would have learned to see an input data point and predict the expected output. Researchers believe supervised learning has become the most widely used machine learning techniques in stock market prediction.

In the literature, different types of machine learning techniques have been developed for utilization in stock market predictions. **Table I** lists few of these techniques along with the literature where they have been

discussed. Due to promising outputs given by the techniques listed below they have become the most frequently applied methods for SMP.

Table I Few techniques of Machine Learning for SMP

Variants of ML Techniques	**Literature where discussed**
Artificial Neural Networks (ANNs)	Nermend & Alsakaa, 2017[40]; O'Connor & Madden, 2006[41]
Support Vector Machines (SVMs)	Cao & Tay, 2001[42]; Huang et al., 2005[43]
Fuzzy Theory	Zadeh, 1965[4]
Fuzzy time-series	Cagcag Yolcu & Alpaslan, 2018[44];Chu et al., 2009[45]
Adaptive Network-based Fuzzy Inference systems	Wei et al., 2011[46]
Random Forests	Khaidem et al., 2016[47];Lohrmann & Luukka, 2018[48]
Decision Trees	Tsai & Hsiao,2010[49]
K-Nearest Neighbour (KNN) classifiers	Zhang et al., 2017[50]

Unlike supervised learning techniques, there is another class of machine learning algorithms termed as unsupervised learning algorithm. The goal of unsupervised learning is to train an algorithm to find a pattern, correlation, or cluster in the given dataset. Moreover, in contrast to the previously discussed supervised learning techniques, the ability of clustering as an unsupervised method was also examined for forecasting stock prices [51].

3. Related Work

The investors and researchers have investigated that stock market prediction is a hot topic of research these days and concerning market forecasting models in finance, literature review studies are comparably rare and only few examples exist.

Kitchenham and Charters [52] concluded a method of presenting a systematic review and discussed that the main objective includes identification, evaluation and discussion of related works for answering the research questions. They also stated that review of literature if left incomplete would result in low scientific value. In addition to that, the plan was proposed for the systematic literature review study, and it was implemented with the process presented in Figure 3.

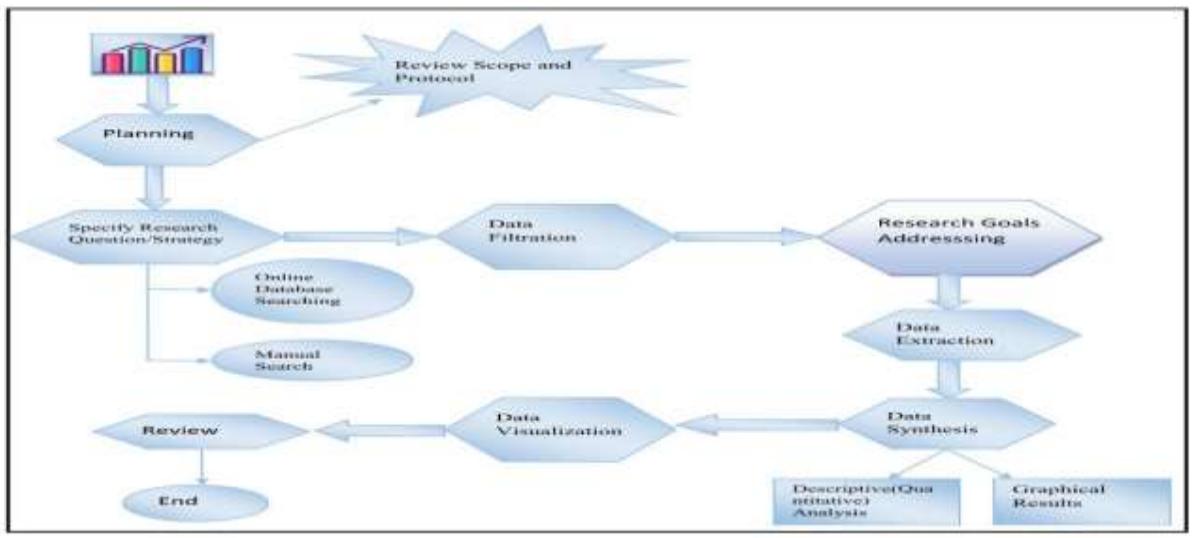

Figure 3. Workflow of Literature review

From many years, financial trading has been the prime interest of many companies with their sole operation being growing exponentially. Along with the companies, market enthusiasts, investors and researchers have extensively researched the area of using machines (machine learning) to predict stock markets. In this article, the main aim is to provide a focused survey on the recent advances of ML techniques for stock market prediction. To collect the relevant research articles for the survey, a systematic survey strategy was carried out and the search was initialised using the keywords "stock market forecasting" or "share market forecasting" along with various ML approaches. Along with various machine learning techniques, their variations have also been adopted for the financial market applications. This research study focuses on two main research questions: (i) Which different computational algorithms are suggested for solving financial problems?; (ii) Which main challenges and research opportunities are still open in this research field?.

Many time-series problems such as video, stock prediction, music and speech recognition were reviewed in L¨ angkvist, Karlsson, and Loutfi [53] by considering unsupervised feature learning.

For building an intelligent trading system, financial markets were studied with computational intelligence in terms of machine learning and NNs. Cavalcante, Brasileiro, Souza, Nobrega, & Oliveira [33] gives an overview of the most important primary studies published from 2009 to 2015, which cover techniques for pre-processing and clustering of financial data, text mining, and other forecasting methods and defined a systematic procedure to build such systems and discussed challenges and open problems.

Li and Ma[54] mentioned few primary studies that use methodology of ANN to prediction of "banking and financial crisis, stock market and exchange rate". No detailed discussion of ANN architectures or learning strategies is being conducted in this study.

Xiao Zhong & David Enke [36] have used the hybrid approach to predict the movement of S&P 500 Electronic Trading Funds. On top of hybrid machine learning algorithms, they have designed classification mining procedures. After pre-processing of raw data, the ANNs and DNNs are used with both the untransformed dataset as well as the PCA represented datasets to forecast the daily market returns.

Nikfarjam, Emadzadeh, and Muthaiyah [55] reviewed database of studies which were implementing text mining techniques for extracting qualitative information of companies. Then they used this information for analysing that news is good or bad for the company and hence modelled stock price predictions according to the news.

An important work related to the present research was proposed by Atsalakis and Valavanis [9] in the year 2009. The authors studied over 100 scientific articles which solved stock market forecasting problem by applying soft computing techniques. They performed five different and important tasks: (i) investigating the financial market, (ii) the input variables used in the market, (iii) benchmarks used in each study (iv)the methods and parameters used to build the predictors and (v) last but not the least the performance measures used for evaluation of the proposed methods. In contrast to it, this research study focuses on investigation of several areas of computational intelligence which are used for stock market forecasting.

One of the latest research study proposed by Tkáč & Verner [56] surveyed those studies which have used ANNs in various business applications, which included auditing and accounting, crediting scoring, financial analysis, inflation, marketing, and many more.

Aguilar-Rivera, Valenzuela-Rendón, and Rodríguez Ortiz [57] review the application of evolutionary computation methods, such as genetic algorithms, learning classifier systems, estimation of distribution algorithms and multi-objective evolutionary algorithms, to solve financial problems.

Cao L. J., Tay E. H. [42] proposed a paper which deals with the application of SVM in financial time series forecasting. The feasibility of applying SVM in financial forecasting is examined by comparing it with the multilayer back-propagation (BP) neural network and the regularized radial basis function (RBF) neural network. The variability in performance of SVM with respect to the free parameters is investigated experimentally. Five real future contracts collected from the Chicago Mercantile Market are used as the data sets. The simulation shows that among the three methods, SVM outperforms the BP neural network in financial forecasting.

Hassan et al [58] proposed the implementation of a model built on a combination of Genetic Algorithms (GA), Artificial Neural Networks and Hidden Markov Model (HMM). The purpose was transforming the daily stock values to independent groups of prices.

Huang et al. [42] proposed an SVM model which predicted financial trend after evaluation of weekly trend of NIKKEI 225 index. The study compared SVM, Linear Discriminant method, Elman Back propagation Neural Networks and Quadratic Discriminant method and found that SVM outperforms others in case of forecasting. The study was based on the fact that an SVM seeks to minimize the upper threshold of the error of its classifications.

Just like every proposed method studies some stock markets as well as indices. Similarly, **Table II** enlists few articles along with the relevant market surveyed.

Table II List of Surveyed stock markets

Reviewed Article	**Reviewed Stock Exchanges**
Chen_&_chen_et_al[6]	Taiwan Stock Exchange
Atsalakis and Valavanis [9]	Athens Stock Exchange, NYSE
Reza et al. [59], Thakur et al. [60] and Wang and Wang [23]	German Stock market(Frankfurt Stock Exchange)
Lu [61], Thakur et al. [59], Dai et al. [62]	Japan Stock market(Tokyo Stock Exchange)
Thakur et al. [59], Patel et al. [63], Pathak and Shetty [64], Senapati et al. [65], Chopra et al. [66]	Indian Stock Market (NSE,BSE)
Asadi et al. [67], Wang and Wang [68], Chien and Chen [69], Niaki and Hoseinzade [70], Hu et al. [71], Wang and Wang [22]	USA stock market(NASDAQ, NYSE)
Dai et al. [61], Wang et al. [72], Pang et al. [73] Zhong et al[36]	China Stock market(Shanghai stock Exchange, Shenzhen stock exchange)
Nayak et al. [74] and Wang and Wang [22]	UK stock market(London stock exchange)

New financial prediction algorithm based on SVM ensemble was proposed by Sun et al. [75]. The method for choosing SVM ensemble's base classifiers was proposed by deeming both diversity analysis and individual prediction. Final results showed that SVM ensemble was importantly better than individual SVM for classification.

Araújo et al. [76] proposed the morphological rank linear forecasting approach to compare its results with time-delay added evolutionary forecasting approach and multilayer perceptron networks. Employing tree-based ensemble methods and deep learning algorithms for predicting the stock and stock market trend is a new area of research activities. In light of employing bagging and majority vote methods, Tsai et al. [49] used two different kinds of ensemble classifiers, such as heterogeneous and homogeneous methods. They also consider macroeconomic features and financial ratios from Taiwan stock market to examine the performance of models. Ballings et al. [77] compared the performance of AdaBoost, Random Forest and kernel factory versus single models involving SVM, KNN, Logistic Regression and ANN. They predict European company's prices for one year ahead. The final results showed that Random Forest outperformed among all models.

Basak et al. [78] employed XGBoost and Random Forest methods for the classification problem to forecast the stock increase or decrease based on previous values. Results showed that the prediction performances have advanced for several companies in comparison with the existing ones. Weng et al. [16] improved four ensemble models, boosting regressor, bagging regressor, neural network ensemble regressor and random forest regressor.

Gandhmal and Kumar [15] performed a systematic analysis and review study using over 50 research articles related to stock market prediction. They essentially classified the selected studies with respect to the applied prediction methods with a detailed discussion of them, the year published, the performance metrics, and the

software tools. Shah et al. [63] also provided a concise review and taxonomy of stock market forecasting models. Bustos and Pomares-Quimbaya [18] presented a systematic review of the prediction methods used in the stock market, covering 52 studies published from 2014 to 2018. This review focused on different types of machine learning techniques, including deep learning, text mining, and ensemble techniques.

Moreover, a study by Jiang[7] surveyed deep learning models applied for stock market predictions in the last three years. It also provided a brief overview of the data used and the data processing methods used and pointed out some future research directions based on existing research. By examining 30 journal and conference articles Kumar et al. [10] provide a complete overview of various aspects adopted in stock market prediction studies, including machine learning algorithms, performance measures, datasets, and journals.

According to Weng et al. [16], the main components in stock market prediction are the tracking of relevant information about data and predictor variables, and the selection of AI techniques that are effective for prediction and analysis. The studies of eminent researchers are listed in the **Table III** which classifies significant works on the basis of machine learning algorithms. From the above research background, it is clear that each of the algorithms can effectively solve stock prediction problems. However, it is vital to notice that there are specific limitations for each of them. The prediction results not only are affected by the representation of the input data but also depend on the prediction method.

Table III

Classification of reviewed articles about financial market prediction using machine learning algorithms

Reviewed Article	Stock Market	Predictive variable	Prediction	Main Learning method	Performance Measure/Results
Huang et al. (2005)[43].	USA, Japan (Chicago Mercantile Exchange, NIKKEI 225 Index)	Currency, Indices	Direction	SVM, Neural networks, LDA	Accuracy(Hit Ratio) LDA 55 EBNN 69 (Elman Backpropagation NN) SVM 73 Combining model 75
Enke and Thawornwong (2005)[79] .	USA	Fundamentalist	Direction	Neural Networks	RMSE (Portfolio class NN-1.0997)
Wang et al. (2012)[23]	China, USA	Prices	Prices	Neural Networks, GA	Proposed Hybrid model with ARIMA+BPNN+EWM Accuracy, MAE, RMSE, MAPE
Kumar and Thenmozhi (2014)[80]	India	Returns	Return	Neural Network, SVM, RF, ARIMA	Proposed ARIMA+SVM ARIMA+RF ARIMA+NN MAE- 0.0121, 0.0124, 0.0123 respectively RMSE- 0.0171,0.0173,0.0172 respectively
Chen et al. (2014)[81]	Taiwan	Indices, TA	Prices	Fuzzy logic	RMSE
Yan et al. (2017)[82]	China(Shanghai Stock	Prices	Prices	Neural Networks(bayesian regularised	MAE 17.0573 MAPE 0.85% MSE 543.6042

	exchange)			NN)	
Weng et al. (2017) [16]	USA	TA, Text	Direction	NNs, SVM, Decision trees	SVM shows good results with Market data +TIs as input Accuracy 0.616 AUC 0.711 F measure 0.633
Barak et al. (2017)[83]	Iran	Fundamentalist	Return and Risk	NNs, SVM, Decision trees	Accuracy
Pan et al. (2018)[27]	USA	Fundamentalist, Prices	Prices	SVM	RMSE, MAE
Patel et al. (2015)[63]	India	TA	Direction	NNs, SVM, RF, NB	Accuracy
Laboissiere et al. (2015)[24]	Brazil	Indices	Maximums, Minimums	NNs	MAE, MAPE, RMSE
Vilela et al.(2019)[51]	Brazil(IBrX index)	Daily Price values	Prices, Returns	K-Means, Fuzzy C-means, SVR	RMSE, MAPE

However, in the literature there are few more studies which also listed techniques for evaluation of working algorithms. These studies did not give exact numerical figures for results and such studies are discussed in **Table IV.**

Table IV Selected Studies for Machine Learning Approach

Study	**Input Data**	**Main approach of Machine Learning**	**Targeted Index**	**Result Measurement Technique**
Hu et. al.(2018)[71]	Google Trends, Indices Data	ANN	S&P 500, DJIA	Hit Ratio
Qiu et. al(2016) [84]	Technical Indicators	ANN	Nikkei 225 index	Hit ratio
Sedighi et. al.(2019)[85]	Technical Indicators	SVM/SVR	DOW 30, NASDAQ 100, S&P 500	RMSE
Zhang, Zhang et. al.(2019)	Closing Prices	Fuzzy Theory	SSECI, TAIEX	RMSE
Lien Minh et. al.(2018)	Financial News, Index Prices	DEEP Learning	S&P 500	Accuracy
Singh and Srivastava(2017)	Technical Indicators	DEEP learning	NASDAQ	RMSE
Weng et. al.(2017)[16]	Market Data, Technical Indicators	Feature Selection	AAPL stock	Accuracy
Zhou et. al.(2018)[27]	Stock Prices and Email Data	Others	Enron stock	Accuracy
Jiang et. al.(2020)[7]	Technical and macroeconomic variables	Stacking method	S&P 500, DOW 30, NASDAQ	Accuracy, Precision, F Score
Carta et. al.(2021)[86]	MLP, Gradient Boosting, RF and other two	Classifier ensembles	S&P 500 index	Accuracy, Precision, F1 score

4. Research Gaps

1. Only 35.3% of the studies analysed profitability, and only two articles implemented risk management.
2. Detail study of articles with statistical forecasting methods is not included.
3. The previous studies which involved analysis of literature primarily focused on either using qualitative or quantitative tools for analysis. The present study is one of its kinds which involved using both the qualitative as well as quantitative measures for analyzing the literature relating to stock returns.
4. Hybrid approaches that combine statistical and machine learning techniques will probably prove to be more useful for stock prediction.
5. Comparative study of Machine learning algorithms and Deep Learning algorithms for Stock Market Prediction seems missing.

5. Methodology and data

5.1 Search Queries

a. Manual search

The initial set of primary studies included nine journal articles (Atsalakis & Valavanis, 2009[9]; Chang & Liu [69]; Chong et al., 2017; Enke & Thawornwong, 2005[79]; Huang et al., 2005[43]; Lohrmann & Luukka, 2019[48]; Patel et al., 2015[63]; Tsai et al., 2011[49]). These articles are found in the manual search for literature on stock market predictions, and decision was taken to use them as the initial set after examining the full text of each of these articles. Moreover, this set of articles was part of the final analysis.

b. Automated Search

The search queries are implemented on online databases such as Scopus, WOS and emphasize the objectives of our research. To define them, a person needs to be experienced and knowledgeable. The terms ‘‘Prediction’’ and ‘‘Forecasting’’ were used separately for each data source, since it was noticed that relevant articles had used at least one of them in their definitions. Prediction or forecasting was expected for either ‘‘Stock price’’ or ‘‘Stock return’’ by using them in the same search query. The term ‘‘Machine Learning’’ was used as the closed term of the defined query since the expected articles are supposed to be based on the forecasting models that applied machine learning methods. However, it became very clear that using ‘‘AI’’ separately with ‘‘Machine Learning’’ resulted in a larger set of relevant articles.

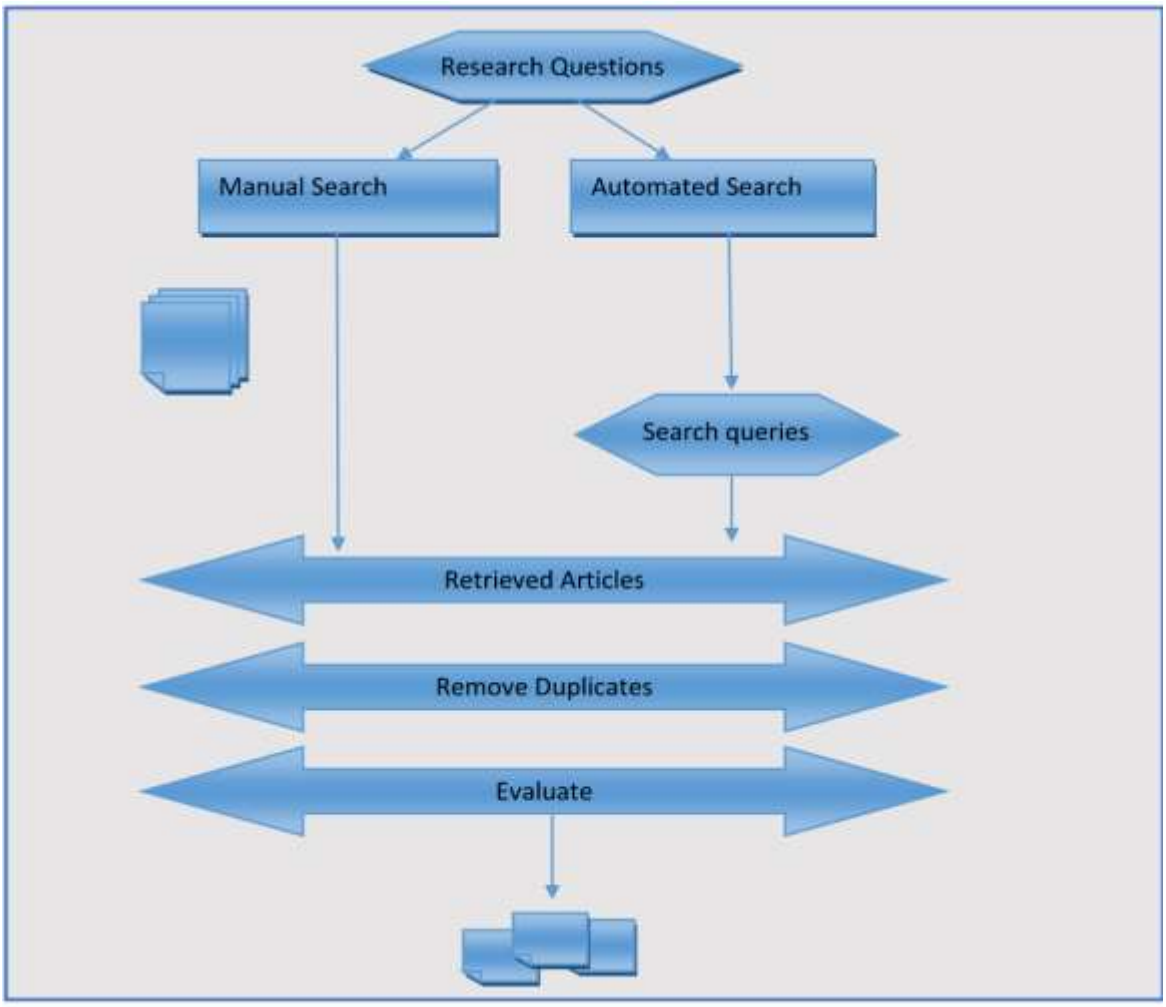

Figure 4 Retrieving mechanisms of data from searched queries

5.2 Data Collection

To ensure the relevance and validity of research, it is important to select high-quality international scientific articles that publish original and reliable sources of information and knowledge. Google Scholars by Google, Web of Science (WOS) by Clarivate Analytics and Scopus by Elsevier are considered the main valuable resources of collecting data required for bibliography information. The Google Scholars is huge database with ample of significant papers, also it can feature any paper from probably any journal [87]. As a trusted and comprehensive bibliographic database, WOS is the most commonly used database in bibliographic research. Choosing the online database named as Scopus, as a source of knowledge can widely obtain high-quality journal articles. The data were searched on 7th April, 2022. The time range is all years. A total of 7484 articles and contributions were categorized relevant according to keywords search strategy. Scopus is the largest abstract and citation database of peer-reviewed literature: scientific journals, books and conference proceedings; delivering a comprehensive overview of the world's research output in the fields of science, technology, medicine, social sciences, and arts and humanities. It also features smart tools to track, analyze and visualize research.
Consequently, as compared to WOS, the decision finalized on Scopus as it's having a good amount of quality journals covering an extensive set of publication [88] and hence it gives a clear picture of the underneath laying research patterns.

5.3 Data selection through search string

Search strategy for reference papers about relevant research on the "Stock market prediction using machine learning is based on the "Advanced Search documents" in the Scopus database. More specifically, we use the keywords (("stock market prediction" OR "Finance market prediction" OR "Financial Market*" OR "share

market prediction" OR "Stock indices prediction" OR "Stock index forecasting" OR "Corporate finance prediction" or "stock indices forecast*") AND ("machine learning" OR "deep learning" OR "artificial intelligence" OR "RNN" OR "LSTM" OR "fuzzy logic")) and search them in the topic. This exercise yields a total of 7377 documents published from 1975–2021.

5.4 Data Extraction

The Comma-separated values (CSV) exported from the Scopus is in a raw format, where individual fields are having multiple values using separators such as ";" and ",". These fields need to be pre-processed to transform as atomic entities and the resulting data is converted into its equivalent tables in Excel database for ease of doing statistical analysis.

So, query outcome before refining processes 7484 documents, but after restricting the language of Research papers to English and filtering the papers of all the other languages, the query outcomes only 7377 Journal articles, conference proceedings, review papers , editorials collectively. The "DOI" is acting as a primary entity to uniquely identify a particular publication. Those publications where there was no "DOI" were eliminated and even the empty rows were deleted in Excel leaving us with 6851 publications in all.

5.5 Data synthesis

Data synthesis is aimed at analyzing and summarizing information observed from the selected articles to answer the research questions posed in this study.

5.6 Data Visualization using Bibliometric Analysis

The quantitative analysis of the research papers, based on any one special topic, through some statistical method is called as Bibliometrics. It is a mathematical way to access the quality of the studies, to find the key findings of the scholars and for the prediction of future studies [2]. There are ample of data sources which are available for bibliography information and the main among them are Google Scholars by Google, Web of Science (WOS) by Clarivate Analytics and Scopus by Elsevier. The Web of Science (WOS), an online database which includes almost all research works and papers along with built in analysis tools and hence produces representative figures. Nowadays, the bibliometric approach of reviewing data has become popular due to various reasons, some of which include the introduction of latest software techniques, interdisciplinary methodologies, and embedding of increased capabilities to handle large volumes of data[89].

Growing financial market is very important for the economic and social organization of modern society. Although a number of methodologies and techniques have been developed for improvement in financial market forecasting, yet the investors can earn significantly higher returns if the prediction accuracy is improved. An ample of research work has been published and carried out in the past to find the suitable forecast models for earning profits in financial markets. Before carrying out the processing in research work, information preparation emerges as a crucial step as it provides a knowledge base as well as helps the analysts and scholars to thoroughly study the work of other experts[88].

Bibliometric method of study of previous research work is considered as scientific as well as statistical method. This method analyses the published research works on the basis of co-authorships, citation relationships, most frequently used keywords, etc. This method has the advantage of using data from various online database sources such as Pubmed, Scopus, and Web of Science. It can also work with various data formats i.e. it can analyse RIS format, CSV format, BIBtex file, text file etc. Bibliographic tools find the associations between the literature for increasing accuracy and convenience for obtaining relevant information. As widely used bibliometric software VOSViewer, it is able to display large bibliometric maps with interpretation of the shown clusters[90]. This software is freely available for researchers on the website www.vosviewer.com. It includes some other advantages like the fact that it allows the detailed study of bibliometric maps constructed from various aspects. The most significant advantage is the colour visualization which helps in identification of individual clusters and also makes this tool user-friendly.

A bibliometric analysis is being followed for the systematic literature overview, incorporating both quantitative and qualitative aspect of literature ([92],[93]). Under this approach, the following analyses are being incorporated: (1) bibliometric citation analysis. (2) Bibliometric co-authorship analysis. (3)

keyword/cartography analysis and (4) bibliographic coupling analysis and (5) content analysis. The Bibliometrix package of R is also used, which includes the graphical interface Biblioshiny [94]. To build visualization networks, VOSViewer will be used to construct visualization networks ([88],[89]); [91] ;[90]. Both programs are famous and the best known and widely used to analyze bibliometric data[95].

6. Results and Discussions

In this study, various types of analysis have been performed and results are concluded. The following analyses have been conducted:

(i) Category Analysis (ii) Citation analysis (iii) Demographic analysis (iv) Keyword analysis
For the visualisation of analysis, VOSViewer is being used. Also few steps of analysis are being performed by BIBLIOMETRIX package of R.

6.1 Category analysis

The Scopus database was used to extract the relevant manuscripts. Consequently, it is taken care of to identify a bibliometric study as in the area of finance.

Figure 5 shows a Pie chart of academic research on stock market prediction. Bigger arc show a higher number of articles. The leading subject is computer science and artificial intelligence (AI) with 5535 papers which are approx 37.1% of total 7377 document results. 2510 papers are in engineering discipline which comes out to be 16.8%. Mathematics contributes 1818 total documents, while business, management and accounting contribute in 733 documents in total. There is an increased interdisciplinary research work found in Scopus database as well as in Web of Science.

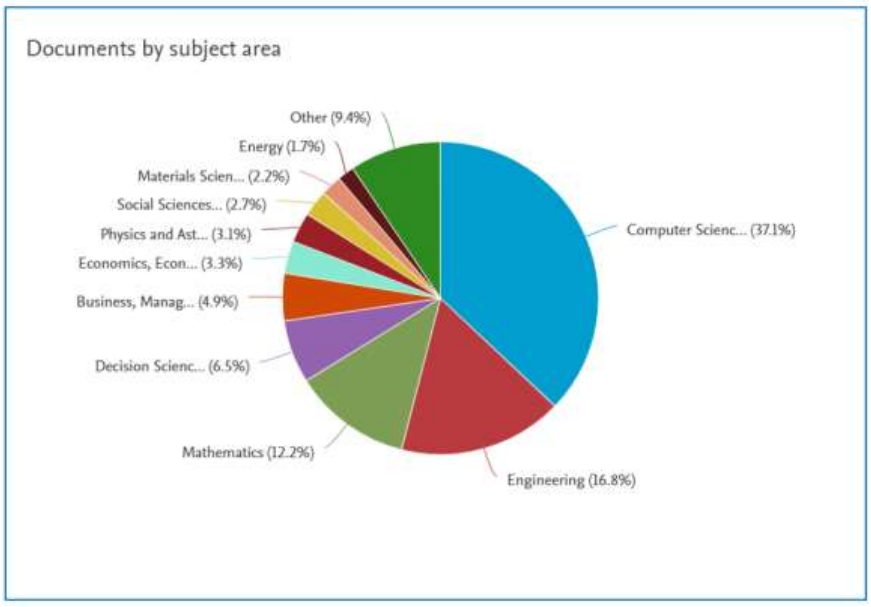

Figure 5 Pie chart showing the results of academic research in SMP.

6.2 Citation based Analysis

This analysis is done through VOSViewer software. It includes citation analysis of authors and co-authors. While carrying out the analysis through VOSviewer, maximum limit of authors per document were limited to only 4, with minimum 4 documents per author to be included in the list and almost every document having minimum citation count of 3. This step yielded 534 authors which satisfied the criteria. Although, all these authors were not connected to each other, hence overall 365 authors were listed which were interconnected to each other. Figure 6 shows the overlay visualisation with different colours. IN the lower left hand of the figure the dimensions of analysing the figure is listed where “Blue” colour represents the authors with least count of citations and “Yellow” colour represents the authors with highest count of citations.

The top most influential author of researches based on stock market forecasting as shown in figure 6 as per VOSViewer metrics until Apr 2022 is recorded as Nath B. with total 5 documents (articles, research papers, conference proceedings etc.) having count of 114.0 average citations in total. This author is followed by Liu C. with total number of 7 documents and having a count of 76.14 average citations in total. **Table 5 and 6** lists the highly cited authors, with number of relevant documents and count of average citations.

Table V Authors with highest average citations listed below in table

Author Name	Total no. of Documents	Average citations
Nath B.	5	114.0
Liu C. –H.	7	76.14
Enke D.	11	57.55
Wang Y.	4	52.50
Chen H.	24	51.25

Table VI Highly cited authors with average citations of their documents

Title	**Reviewed Article**	**Total count of Citations**
Kara et. al.(2011)[96]	"Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul stock exchange,"	437
Kayakutlu et. al.(2011)[22]	"Using artificial neural network models in stock market index prediction"	439
Leung et. al.(2000)[97]	"Forecasting stock indices: A comparison of classification and level estimation models"	229
Zhang & Wu(2009) [98]	"Stock market prediction of S&P 500 via combination of improved BCO approach and BP neural network"	275
Boyacioglu & Avci(2010)[30]	"An adaptive network-based fuzzy inference system (ANFIS) for the prediction of stock market return: The case of the Istanbul stock exchange"	202

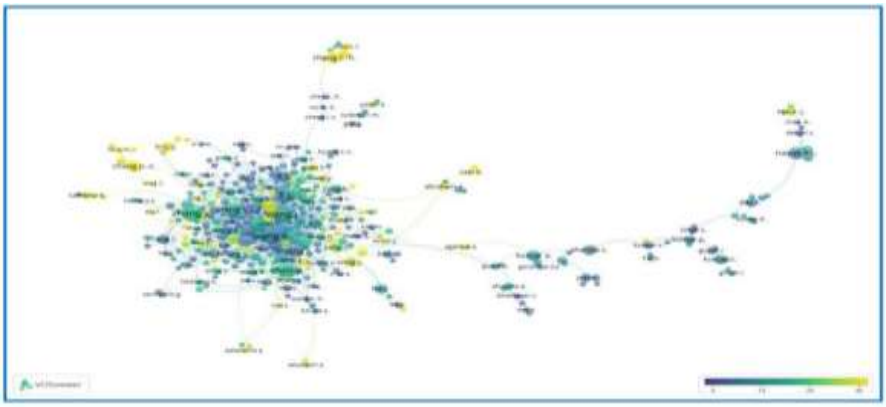

Figure 6. Co-authorship analysis conducted using VOSViewer.

Figure 7 shows the leading journals publishing bibliometric review of Stock market prediction. This Figure compares the document count for up to 5 sources which have published highest number of contents from the year 1992 till 2022.

IEEE Access has the highest number of documents published in the year 2020 and the count of total documents published in IEEE Access comes out to be 306. Second position is bagged by the journal Advances in Intelligent Systems and computing with 285 documents out of the whole list of published research materials.

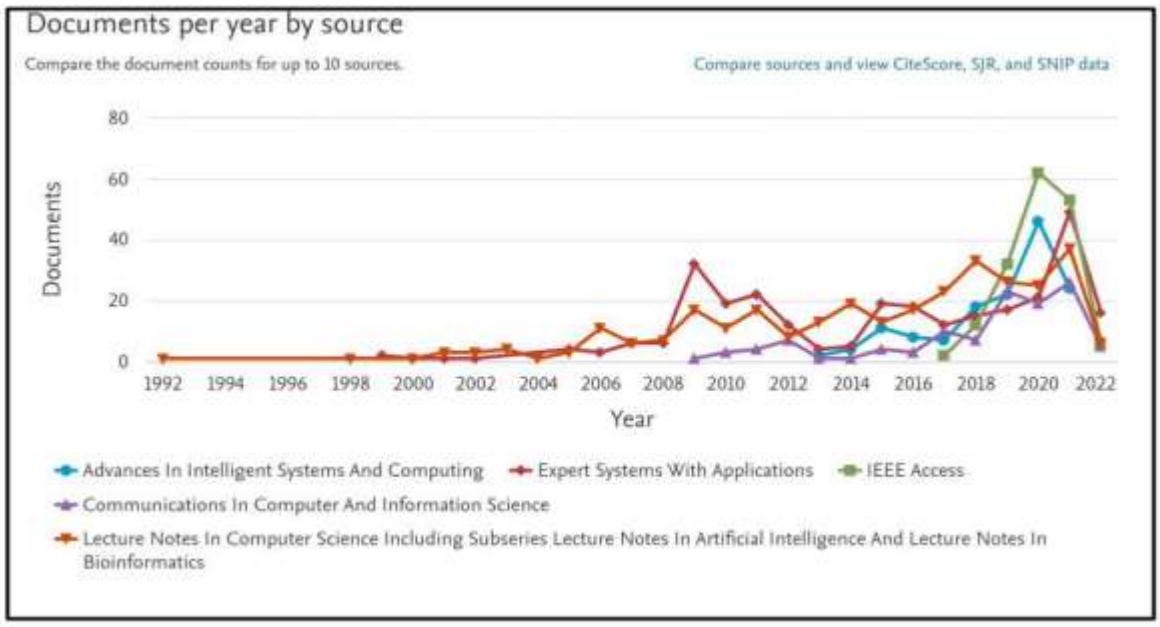

Figure 7 Journals showing the maximum count of publications in SMP.

6.3 Demographic Analysis

The number of published papers and global citations about a research by a country can reflect the country's scientific research level and emphasis. This analysis compares the document count of top 15 countries which have been involved in the search related to machine learning in stock market prediction. In the figure 8 it is clearly depicted that China has published a lot of Research work on Stock market prediction with 1766 as total count of documents out of 7377 documents. Afterwards India and United States have published the contents related to this research with 1302 and 812 total research materials.

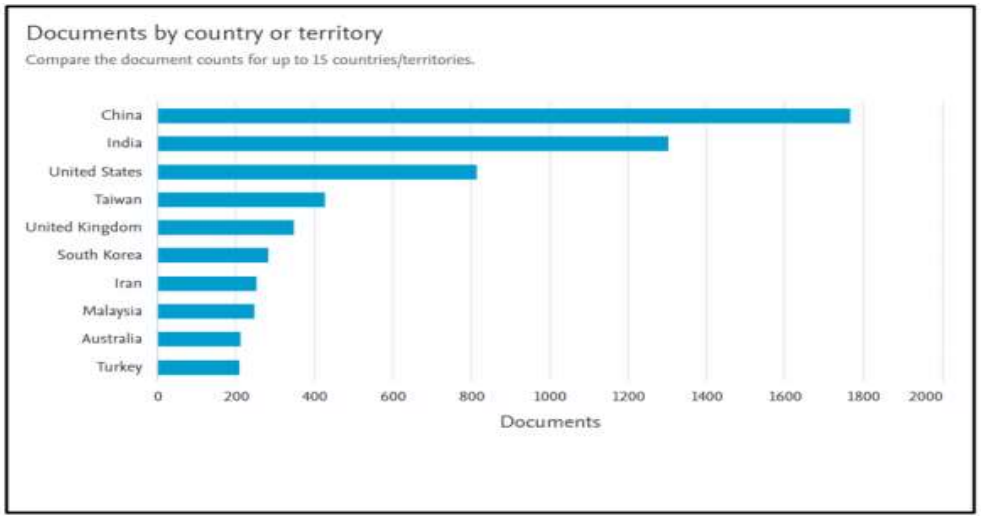

Figure 8 Top 15 Countries showing the related research

6.4 Keyword/cartographic analysis

This research conducted keyword analysis to conceptualize the evolution and dynamics of bibliometric studies in finance literature [90]. To have meaningful analysis, filtration is required for a minimum threshold of two for the occurrence of a particular keyword to be included in the analysis (yielding a total of 2674 keywords) [99], and exclude non-related and method-based terms. Ultimately, through VOSViewer analysis 1987 keywords with total link strength of 1721, is reported in Figure 9. Since the frequency of keywords is less before 2018, the time span is limited to three years up to the early start of 2022. It is found that the most repeated keywords in bibliometric reviews in finance are 'forecasting,' 'commerce,' 'machine learning,' 'finance,' and 'stock.' The association strength normalization method used by VOS Viewer identified 5 clusters in five different colours, where red coloured keywords represent terms related to forecasting such as "vector machine", "forecasting model", "stock index", "financial time series" etc. Yellow colour represented keywords related to machine learning methods like "HMM", "lstm", "CNN" , "recognition", "Detection" etc. Green colour represented ample keywords like "sentiment", "mining", "intelligence" etc. The other two clusters for example, blue and purple are relatively small in size. Hence, it can be concluded that two most important clusters are red and yellow which

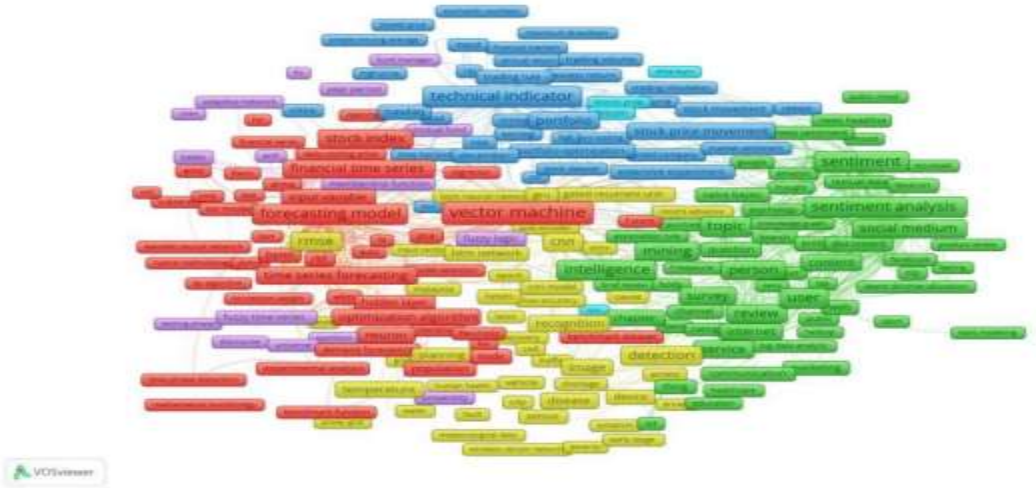

Figure 9 Co-occurrences of keyword using VOSViewer software found in the final list

Next analysis is conducted using Biblioshiny () function in Bibliometrix package. This library helps to import data from online databases and then helps to perform bibliographic coupling, co-occurrence of keywords/ authors in the final list of documents.

Figure 10 The word cloud of the relevant keywords found.

7. Conclusion

With the inspiration from rapid development and increased usage of machine learning models for stock market prediction, this article reviews the recent progress after surveying around 100 related published articles. Each step from raw data collection and data processing to study of frequently used machine learning based prediction models are covered in the literature. The article showed the detailed presentation of existing forecasting approaches based on machine learning and where they have been used in the literature. It also analyzes the stocks and financial markets as well as the types of approaches used for predictions. The review also conducted the bibliometric analysis of this study of machine learning in stock market forecasting. The techniques utilized for the stock market prediction involves ANN, SVM, SVR, HMM, NN, fuzzy based techniques, K-means, and so on. Out of them, commonly used technique for attaining effective stock market prediction is ANN and the fuzzy-based technique. In the future, these techniques as well as hybrid techniques by merging one or few of them can be effectively utilized for monitoring the stock market. The major challenge faced by the stock price prediction systems is that accuracy cannot be obtained by only using historical stock data as they are affected due to various factors, which involve government policy decisions, market sentiments and so on. Also, the data pre-processing is the most complex task for data mining. These are major limitations that need to be addressed in the future by adapting advanced stock market prediction techniques.

References

[1] Nassirtoussi, A.K. and Aghabozorgi, S. et. al. (2014) Text mining for market prediction: A systematic review. *Expert Syst Appl.*, 41(16),7653-7670,10.1016/j.eswa.2014.06.009

[2] Göçken, M. and Özçalici, M. and Boru, A. and Dosdotru, A. T. (2015) Integrating Metaheuristics and Artificial Neural Networks for improved Stock Price Prediction. *Expert Systems with Appl.*, 10.1016/j.eswa.2015.09.029

[3] Fama, E. F. (1995) Fama_RandomWalksStockPrices. *Financial Analysts Journal*, 51(1),75–80.

[4] Thakkar, A. and Chaudhari, K. (2021) Fusion in stock market prediction: A decade survey on the necessity, recent developments, and potential future directions. *Inf. Fusion.*, 65, 95–107, 10.1016/j.inffus.2020.08.019

[5] Thakkar, A. and Chaudhari, K. (2020) CREST: Cross-Reference to Exchange-based Stock Trend Prediction using Long Short-Term Memory. *Procedia Comput. Science.*, 167, 616–625, 10.1016/j.procs.2020.03.328
[6] Chen, Y. J. and Chen, Y. M. and Tsao, S. T. and Hsieh, S. F. (2018) A novel technical analysis-based method for stock market forecasting. *Soft Comput.*, 22(4),1295–1312, 10.1007/s00500-016-2417-2
[7] Jiang, W. (2021) Applications of deep learning in stock market prediction: Recent progress. *Expert System Applications.*, 184, 115537, 10.1016/j.eswa.2021.115537.
[8] Keller, C. and Siegrist, M. (2006) Investing in stocks: The influence of financial risk attitude and values-related money and stock market attitudes. *J Econ Psychol.*, 27(2), 285–303, 10.1016/J.JOEP.2005.07.002
[9] Atsalakis, G. S. and Valavanis, K. P. (2009) Surveying stock market forecasting techniques - Part II: Soft computing methods. *Expert Syst. Appl.*, 36(3)-II, 5932–5941, 10.1016/j.eswa.2008.07.006.
[10] Kumar, G. and Jain, S. and Singh, U. P. (2021) Stock Market Forecasting Using Computational Intelligence: A Survey. *Springer Netherlands*. 28(3).
[11] Henrique, B. M. and Sobreiro, V. M. and Kimura, H. (2019) Literature review: Machine learning techniques applied to financial market prediction. *Expert Syst. Appl.*, 124, 226–251, 10.1016/j.eswa.2019.01.012.
[12] Fama, E. F. (1991) Efficient Capital Markets: II. *J. Finance.*, 46(5), 1575, 10.2307/2328565.
[13] Farias Nazário, R. T. and e Silva, J. L. and Sobreiro, V. A. and Kimura, H. (2017) A literature review of technical analysis on stock markets. *Q. Rev. Econ. Financ.*, 66, 115–126, 10.1016/j.qref.2017.01.014.
[14] Kumbure, M. M. and Lohrmann, C and Luukka, P. and Porras, J. (2021) Machine learning techniques and data for stock market forecasting: A literature review. *Expert Syst. Appl.*, 197, 116659. 10.1016/j.eswa.2022.116659.
[15] Gandhmal, D. P. and Kumar, K. (2019) Systematic analysis and review of stock market prediction techniques. *Comput. Sci. Rev.*, 34, 100190, 10.1016/j.cosrev.2019.08.001.

[16] Weng, B. and Ahmed, M. A. and Megahed, F. M. (2017) Stock market one-day ahead movement prediction using disparate data sources. *Expert Syst. Appl.*, 79, 153–163, 10.1016/j.eswa.2017.02.041.
[17] Li, A. W. and Bastos, G. S. (2020) Stock market forecasting using deep learning and technical analysis: A systematic review. *IEEE Access*., 8, 185232–185242, 10.1109/ACCESS.2020.3030226.
[18] Bustos, O. and Pomares, A-Quimbaya (2020) Stock market movement forecast: A Systematic review. *Expert Syst. Appl.*, 156, 113464, 10.1016/j.eswa.2020.113464.
[19] Vanstone, B. and Finnie, G. and Hahn, T. (2012) Creating trading systems with fundamental variables and neural networks: The Aby case study. *Math. Comput. Simul*., 86, 78–91, 10.1016/j.matcom.2011.01.002.
[20] Wei, L. Y. and Chen, T. L. and Ho, T. H. (2011) A hybrid model based on adaptive-network-based fuzzy inference system to forecast Taiwan stock market. *Expert Syst. Appl.*, 38(11), 13625–13631, 10.1016/j.eswa.2011.04.127.
[21] Adio, T. A. and Arogundade, O. (2005) Translated Nigeria Stock Market Prices Using Artificial Neural Network for Effective Prediction. *J. Theor. Appl. Inf. Technol.*, 9(1).
[22] Guresen, E. and Kayakutlu, G. and Daim, T. U. (2011) Using artificial neural network models in stock market index prediction. *Expert Syst. Appl.*, 38(8), 10389–10397, 10.1016/j.eswa.2011.02.068.
[23] Wang, J. J. and Wang, J. Z. and Zhang, Z. G. and Guo, S. P. (2012) Stock index forecasting based on a hybrid model. *Omega.*, 40(6), 758–766, 10.1016/j.omega.2011.07.008.
[24] Laboissiere, L. A. and Fernandes, R. A. S. and Lage, G. G. (2015) Maximum and minimum stock price forecasting of Brazilian power distribution companies based on artificial neural networks. *Appl. Soft Comput. J.*, 35, 66–74, 10.1016/j.asoc.2015.06.005.
[25] Rahman, A D A A F A M T (2017) Predicting Stock Trends Using Tsk-fuzzy Rule Based System. J. Energy Nat. Resour. Manag. , 4(1), 48–55, 10.26796/jenrm.v4i1.74.
[26] Thanh, D. and Minh Hai, N. and Hieu, D. D. (2018) Building unconditional forecast model of Stock Market Indexes using combined leading indicators and principal components: application to Vietnamese Stock Market. *Indian J. Sci. Technology*., 11(2), 1–13, 10.17485/ijst/2018/v11i2/104908.
[27] Zhou, X and Pan, Z and Hu, G et. al.(2018) Stock Market Prediction on High-Frequency Data Using Generative Adversarial Nets. *Math. Probl. Eng.*, 2018, 10.1155/2018/4907423.
[28] Umoru, D. and Nwokoye, G. A. (2018) FAVAR Analysis of Foreign Investment with Capital Market Predictors: Evidence on Nigerian and Selected African Stock Exchanges. *Acad. J. Econ. Stud.*, 4(1), 12–20.

[29] Haleh, H. and Moghaddam, B. A. and Ebrahimijam S (2011) A New Approach to Forecasting Stock Price with EKF Data Fusion. *Int. J. Trade, Econ. Financ*., 2(2), 109–114, 10.7763/ijtef.2011.v2.87.
[30] Boyacioglu, M. A. and Avci (2010) 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), 7908–7912, 10.1016/j.eswa.2010.04.045.
[31] Pérez-Cruz, F. and Afonso-Rodríguez, J. A. and Giner, J. (2003) Estimating GARCH models using support vector machines. *Quant. Financ.*, 3(3), 163–172, 10.1088/1469-7688/3/3/302.
[32] Ravi Kumar, P. and Ravi, V. (2007) Bankruptcy prediction in banks and firms via statistical and intelligent techniques - A review. *Eur. J. Oper. Res*., 180(1), 1–28, 10.1016/j.ejor.2006.08.043.
[33] Cavalcante, R. C. and Brasileiro, R. C. et. al. (2016) Computational Intelligence and Financial Markets: A Survey and Future Directions. *Expert Syst. Appl*., 55, 194–211, 10.1016/j.eswa.2016.02.006.
[34] Chen, H. and Xiao, K. et. al. (2017) A double-layer neural network framework for high-frequency forecasting. *ACM Trans. Manag. Inf. Syst.* ,7(4), 10.1145/3021380.
[35] Mukherjee, S. and Sadhukhan, B. et. al. (2021) Stock market prediction using deep learning algorithms. *CAAI Trans. Intell. Technol*., 10.1049/cit2.12059.
[36] Zhong, X. and Enke, D. (2017) Forecasting daily stock market return using dimensionality reduction. *Expert Syst. Appl.*, 67, 126–139, 10.1016/j.eswa.2016.09.027.
[37] Hiransha, M. and Gopalakrishnan, E. A. et. al. (2018) NSE Stock Market Prediction Using Deep-Learning Models, *ICCIDS 2018.*, Coimbatore, India, pp. 1351–1362, Procedia Computer Science, j.procs.2018.05.050.
[38] Velay, M. and Daniel, F. (2018) Stock Chart Pattern recognition with Deep Learning [Online]. *Available: http://arxiv.org/abs/1808.00418.*
[39] Kubat, M. (2017) An *Introduction to Machine Learning*. , Springer Cham, Coral Gables, USA.
[40] Nermend, Y. and Alsakaa, K. (2017) Back-propagation artificial neural networks in stock market forecasting -An application to the Warsaw Stock Exchange. *Ieb Int. J. Financ*., 15, 88–99, 10.5605/IEB.15.5.
[41] O'Connor N and Madden M G (2006) A neural network approach to predicting stock exchange movements using external factors. Knowledge-Based Syst. 19(5): 371–378. https://doi.org/10.1016/j.knosys.2005.11.015.
[42] Cao L J and TayF E H (2003) Support vector machine with adaptive parameters in financial time series forecasting. IEEE Trans. Neural Networks. 14(6):1506–1518 https://doi.org/10.1109/TNN.2003.820556.
[43] Huang W, Nakamori Y and Wang S Y (2005) Forecasting stock market movement direction with support vector machine. Comput. Oper. Res. 32(10):2513–2522 https://doi.org/10.1016/j.cor.2004.03.016.
[44] Cagcag Yolcu O and Lam H K (2017) A combined robust fuzzy time series method for prediction of time series. Neurocomputing. 247:87–101. https://doi.org/10.1016/j.neucom.2017.03.037.
[45] Wong W K, Xia M and Chu W C (2010) Adaptive neural network model for time-series forecasting. Eur. J. Oper. Res. 207(2): 807–816. https://doi.org/10.1016/j.ejor.2010.05.022.
[46] Wei L Y, Chen T L and Ho T H (2011) A hybrid model based on adaptive-network-based fuzzy inference system to forecast Taiwan stock market. Expert Syst. Appl. 38(11):13625–13631. https://doi.org/10.1016/J.ESWA.2011.04.127.
[47] Khaidem L, Saha S and Dey S R (2016) Predicting the direction of stock market prices using random forest. 00(00):1–20. http://arxiv.org/abs/1605.00003.
[48] Lohrmann C and Luukka P (2019) Classification of intraday S&P500 returns with a Random Forest. Int. J. Forecast.. 35(1): 390–407. https://doi.org/10.1016/j.ijforecast.2018.08.004.
[49] Tsai C F and Hsiao Y C (2010) Combining multiple feature selection methods for stock prediction: Union, intersection, and multi-intersection approaches. Decis. Support Syst. 50(1): 258–269. https://doi.org/10.1016/j.dss.2010.08.028.
[50] Zhang N, Lin A and Shang P (2017) Multidimensional k-nearest neighbor model based on EEMD for financial time series forecasting. Phys. A Stat. Mech. its Appl. 477:161–173. https://doi.org/10.1016/j.physa.2017.02.072.
[51] Vilela L F S, Leme R C, Pinheiro C A M and Carpinteiro O A S (2019) Forecasting financial series using clustering methods and support vector regression. Artif. Intell. Rev. 52(2): 743–773. https://doi.org/10.1007/s10462-018-9663-x.
[52] Kitchenham B A and Charters S (2007) Guidelines for performing Systematic Literature Reviews in

Software Engineering. EBSE Technical Report EBSE-2007-01. School of Computer Science and Mathematics, Keele University. Pp. 1–57.
[53] Längkvist M, Karlsson L and Loutfi A (2014) A review of unsupervised feature learning and deep learning for time-series modeling. Pattern Recognit. Lett. 42(1):11–24. https://doi.org/10.1016/j.patrec.2014.01.008.
[54] Li Y and Ma W (2010) Applications of artificial neural networks in financial economics: A survey. Proc. - 2010 Int. Symp. Comput. Intell. Des. Isc. 2010. 1:211–214. https://doi.org/10.1109/ISCID.2010.70.
[55] Nikfarjam A, Emadzadeh E and Muthaiyah S (2010) Text mining approaches for stock market prediction. 2010 2nd Int. Conf. Comput. Autom. Eng. ICCAE 2010. 4:256–260. https://doi.org/10.1109/ICCAE.2010.5451705.
[56] Tkáč M and Verner R (2016) Artificial neural networks in business: Two decades of research. Applied Soft computing. http://dx.doi.org/doi:10.1016/j.asoc.2015.09.040.
[57] Aguilar-Rivera R, Valenzuela M-Rendón, and Rodríguez J J-Ortiz (2015) Genetic algorithms and Darwinian approaches in financial applications: A survey. Expert Syst. Appl. 42(21):7684–7697. https://doi.org/10.1016/j.eswa.2015.06.001.
[58] Hassan M R, Nath B and Kirley M (2007) A fusion model of HMM, ANN and GA for stock market forecasting. Expert Syst. Appl. 33(1):171–180. https://doi.org/10.1016/j.eswa.2006.04.007.
[59] Hafezi R, Shahrabi J and Hadavandi E (2015) A bat-neural network multi-agent system (BNNMAS) for stock price prediction: Case study of DAX stock price. Appl. Soft Comput. 29:196–210. https://doi.org/10.1016/j.asoc.2014.12.028.
[60] Kumar D, Meghwani S S and Thakur M (2016) Proximal support vector machine based hybrid prediction models for trend forecasting in financial markets. J. Comput. Sci. 17:1–13. https://doi.org/10.1016/j.jocs.2016.07.006.
[61] Chen Y J, Chen Y M and Lu C (2017) Enhancement of stock market forecasting using an improved fundamental analysis-based approach Soft Comput. 21(13):3735–3757. https://doi.org/10.1007/s00500-016-2028-y.
[62] Dai W, Wu J Y and Lu C J (2012) Combining nonlinear independent component analysis and neural network for the prediction of Asian stock market indexes. Expert Syst. Appl. 39(4): 4444–4452. https://doi.org/10.1016/j.eswa.2011.09.145.
[63] Patel J, Shah S, Thakkar P and Kotecha K (2015) Predicting stock market index using fusion of machine learning techniques. Expert Syst. Appl. 42(4): 2162–2172. https://doi.org/10.1016/j.eswa.2014.10.031.
[64] Pathak A and Shetty N P (2019) Indian Stock Market Prediction Using Machine Learning and Sentiment Analysis. Computational Intelligence in Data Mining, Advances in Intelligent Systems and Computing. 711. https://doi.org/10.1007/978-981-10-8055-5_53
[65] Senapati M R, Das S and Mishra S (2018) A Novel Model for Stock Price Prediction Using Hybrid Neural Network. J. Inst. Eng. Ser. B. 99(6):555–563. https://doi.org/10.1007/s40031-018-0343-7.
[66] Chopra S, Yadav D and Chopra A N (2019) Artificial Neural Networks Based Indian Stock Market Price Prediction : Before and After Demonetization. Int. J. Swarm Intell. Evol. Comput. 8(1):1–7. www.researchgate.net/publication/331114490.
[67] Asadi S, Hadavandi E, Mehmanpazir F and Nakhostin M M (2012) Hybridization of evolutionary Levenberg-Marquardt neural networks and data pre-processing for stock market prediction. Knowledge-Based Syst. 35:245–258. https://doi.org/10.1016/j.knosys.2012.05.003.
[68] Wang J and Wang J (2015) Forecasting stock market indexes using principle component analysis and stochastic time effective neural networks. Neurocomputing. 156:68–78. https://doi.org/10.1016/j.neucom.2014.12.084.
[69] Chang Chien Y W and Chen Y L (2010) Mining associative classification rules with stock trading data-A GA-based method. Knowledge-Based Syst. 23(6): 605–614. https://doi.org/10.1016/j.knosys.2010.04.007.
[70] Niaki S T A and Hoseinzade S (2013) Forecasting S&P 500 index using artificial neural networks and design of experiments. J. Ind. Eng. Int. 9(1):1–9. https://doi.org/10.1186/2251-712X-9-1.
[71] Hu H, Tang L, Zhang S and Wang H (2018) Predicting the direction of stock markets using optimized neural networks with Google Trends. Neurocomputing. 285:188–195. https://doi.org/10.1016/j.neucom.2018.01.038.
[72] Wang J, Wang J, Fang W and Niu H (2016) Elman Recurrent Random Neural Networks. Hindawi Publ.

Corp. Comput. Intell. Neurosci. 2016:1–14. http://dx.doi.org/10.1155/2016/4742515.
[73] Pang X, Zhou Y, Wang P, Lin W and Chang V (2020) An innovative neural network approach for stock market prediction. J. Supercomput. 76(3):2098–2118. https://doi.org/10.1007/s11227-017-2228-y.
[74] Nayak S C, Misra B B and Behera H S (2017) Artificial chemical reaction optimization of neural networks for efficient prediction of stock market indices. Ain Shams Eng. J. 8(3):371–390. https://doi.org/0.1016/j.asej.2015.07.015.
[75] Sun J and Li H (2012) Financial distress prediction using support vector machines: Ensemble vs. individual. Appl. Soft Comput. J. 12(8):2254–2265. https://doi.org/10.1016/j.asoc.2012.03.028.
[76] Araújo R D A and Ferreira T A E (2013) A Morphological-Rank-Linear evolutionary method stock market prediction. Inf. Sci. (Ny). 237: 3–17. https://doi.org/10.1016/j.ins.2009.07.007.
[77] Ballings M, Van Den Poel D, Hespeels N and Gryp R (2015) Evaluating multiple classifiers for stock price direction prediction. Expert Syst. Appl. 42(20): 7046–7056. https://doi.org/10.1016/j.eswa.2015.05.013.
[78] Basak S, Kar S, Saha S, Khaidem L and Dey S R (2017) Predicting the direction of stock market prices using tree-based classifiers. North Am. J. Econ. Financ. 47(2017):552–567. https://doi.org/10.1016/j.najef.2018.06.013.
[79] Thawornwong S, Enke D and Dagli C (2003) Neural networks as a decision maker for stock trading: A technical analysis approach. Int. J. Smart Eng. Syst. Des. 5(4):313–325. https://doi.org/10.1080/10255810390245627.
[80] Kumar M and Thenmozhi M (2014) Forecasting stock index returns using ARIMA-SVM, ARIMA-ANN, and ARIMA-random forest hybrid models. Int. J. Banking, Account. Financ. 5(3): 284–308. https://doi.org/10.1504/IJBAAF.2014.064307.
[81] Chen Y S, Cheng C H and Tsai W L (2014) Modeling fitting-function-based fuzzy time series patterns for evolving stock index forecasting. Appl. Intell. 41(2):327–347. https://doi.org/10.1007/s10489-014-0520-6.
[82] Yan D, Zhou Q, Wang J and Zhang N (2017) Bayesian regularisation neural network based on artificial intelligence optimisation. Int. J. Prod. Res. 55(8):2266–2287. https://doi.org/10.1080/00207543.2016.1237785.
[83] Barak S, Arjmand A and Ortobelli O (2017) Fusion of multiple diverse predictors in stock market. Inf. Fusion. 36:90–102. https://doi.org/10.1016/j.inffus.2016.11.006.
[84] Qiu M and Song Y (2016) Predicting the direction of stock market index movement using an optimized artificial neural network model. PLoS One. 11(5):1–11. https://doi.org/10.1371/journal.pone.0155133.
[85] Sedighi M, Jahangirnia H, Gharakhani M and Fard S F (2019) A novel hybrid model for stock price forecasting based on metaheuristics and support vector machine. Data. 4(2):1–28. https://doi.org/10.3390/data4020075.
[86] Carta S, Corriga A, Ferreira A, Podda A S and Recupero D R (2021) A multi-layer and multi-ensemble stock trader using deep learning and deep reinforcement learning. Appl. Intell. 51(2):889–905. https://doi.org/10.1007/s10489-020-01839-5.
[87] Sood S K, Kumar N and Saini M (2021) Scientometric analysis of literature on distributed vehicular networks : VOSViewer visualization techniques. Springer Netherlands. 54(8).
[88] Janková Z (2021) A bibliometric analysis of artificial intelligence technique in financial market. Sci. Pap. Univ. Pardubice, Ser. D Fac. Econ. Adm. 29(3). https://doi.org/10.46585/sp29031268.
[89] Donthu N, Kumar S, Mukherjee D, Pandey N and Lim W M (2021) How to conduct a bibliometric analysis: An overview and guidelines. J. Bus. Res. 133:285–296. https://doi.org/10.1016/j.jbusres.2021.04.070.
[90] Van Eck N J and Waltman L (2010) Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics. 84(2):523–538. https://doi.org/10.1007/s11192-009-0146-3.
[91] Bahoo S (2020) Corruption in banks: A bibliometric review and agenda. Financ Res Lett. 35. https://doi.org/10.1016/j.frl.2020.101499.
[92] Baker H K, Kumar S and Pattnaik D (2020) Twenty-five years of Review of Financial Economics: A bibliometric overview. Rev Financ Econ. 38(1):3–23. https://doi.org/10.1002/rfe.1095.
[93] Baker H K, Kumar S and Pattnaik D (2020) Fifty years of The Financial Review: A bibliometric overview. Financ Rev. 55(1):7–24. https://doi.org/10.1111/fire.12228.
[94] Aria M and Cuccurullo C (2017) Bibliometrix: An R-tool for comprehensive science mapping analysis. J. Informetr. 11(4):959–975. https://doi.org/10.1016/j.joi.2017.08.007.
[95] Khan A, Goodell A W, Hassan M K and Paltrinieri A (2021) A bibliometric review of finance bibliometric

papers. Financ Res Lett. https://doi.org/10.1016/j.frl.2021.102520.
[96] Kara Y, Acar Boyacioglu M, and Baykan O K (2011) Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange. Expert Syst Appl. 38(5):5311–5319. https://doi.org/10.1016/j.eswa.2010.10.027.
[97] Leung M T, Daouk H and Chen A S (2000) Forecasting stock indices: A comparison of classification and level estimation models. Int J Forecast. 16(2):173–190. https://doi.org/10.1016/S0169-2070(99)00048-5.
[98] Yudong Z and Lenan W (2009) Stock market prediction of S&P 500 via combination of improved BCO approach and BP neural network. Expert Syst Appl. 36(5):8849–8854. https://doi.org/10.1016/j.eswa.2008.11.028.
[99] Khan A, Hassan M K, Paltrinieri A, Dreassi A, and Bahoo S (2020) A bibliometric review of takaful literature. Int Rev Econ Financ. 69:389–405. https://doi.org/10.1016/j.iref.2020.05.013.
[100] Dutta, Avijan, Gautam Bandopadhyay, and Suchismita Sengupta. Prediction of Stock Performance in Indian Stock Market Using Logistic Regression. International Journal of Business and Information. 7:105–36.