Multidisciplinary Reviews

REVIEW ARTICLE
Published Online: January 15, 2024
https://doi.org/10.31893/multirev.2024060

Stock market and securities index prediction using artificial intelligence: A systematic review

Harmanjeet Singh[a] ✉ | **Manisha Malhotra**[a]

[a]University Institute of Computing, Chandigarh University, SAS Nagar, Mohali, 140413, Punjab-India.

Abstract The recognition of the value and importance of recognizing patterns in the stock market is widely accepted. As a result, using innovative decision-making strategies is expected to lead to significant returns in stock prices. Predicting the stock market proves challenging because of the limited volatility and inherent disorderliness observed within the data. Hence, investors need help optimizing their profits through informed predictions in the stock market. Stock market projections are made by employing mathematical approaches and machine learning methods. The stock market is inherently vulnerable to unforeseen changes due to its intricate nature and absence of a linear progression. Since the advent of personal computers and the proliferation of technological advancements in the 1990s, scholars have investigated the application of artificial intelligence in the investment sector. Many solutions have been created to tackle the matter of volatility in stock market prices. This study examined a total of 146 research articles that were published in academic journals over twelve years (2011-2023). These papers focused on applying artificial intelligence in predicting stock market trends. The listed works encompass several methodologies, including technical analysis, fundamental analysis, sentiment analysis, and time series analysis. Every academic field is comprehensively explored, encompassing its initial findings and most recent advancements. Moreover, the existing body of literature suggests a growing focus on this particular sector, with a heightened level of specialization and an expanded range of topics being explored.

Keywords: artificial intelligence, finance forecasting, sentiment analysis, share market, trade forecast

1. Introduction

Artificial intelligence in share market investments has been extensively studied since the 1990s when computational methods for financial research were established (Lin, 2019; Mehtab, 2020; Gao, 2018). Computer-aided financial investment automation eliminates "momentary irrationality," finds patterns humans miss, and consumes real-time information. Computational finance describes this understanding. AI has become increasingly crucial in computational finance. 90\% of hedge fund trading is still done manually, despite machines handling most trades (Cheng et al., 2019). Artificial intelligence's growing application has space for growth. AI can optimize financial portfolios, predict asset values, and analyze news and social media sentiment. Some research has recommended combining methodologies from various fields despite their differences. Computational finance also studies investor behavioral analysis, network analysis, financial asset clustering, and dynamic system control in the financial market. This study investigates each issue from its inception to its latest achievements. A sample was collected to compare and examine share price prediction with artificial intelligence works from 2011 to 2023. The subsequent passage provides a systematic review of the paper's contribution. Initially, our investigation will identify and assess potential rationales for the prevailing stock price, employing methodologies such as fundamental and technical research. In the second part of the study, scholars in stock price forecasting undertake a comprehensive analysis of the various techniques, frameworks, data sets, and machine learning models developed for the specific objective of predicting stock prices. Several proposals have been recently made for deploying and enhancing algorithms to predict the market price of stock securities. This systematic review aims to synthesize and categorize many papers about this subject matter based on distinct prediction approaches to elucidate these advancements. The authors have curated a collection of 146 scholarly works that have lately achieved noteworthy advancements in the domain of stock markets. The subject matter of the stock market is extensive, and the articles provided encompass a wide range of these areas. All of the works mentioned above are publications that have been released in the recent past. The articles that present algorithms and data employed in their research are categorized into separate areas based on their target audience.

This comprehensive examination of the prominent stock market prediction techniques and their applications can be a first reference for aspiring researchers in this domain. In contrast to alternative surveys, the present study comprehensively categorizes the diverse methodologies employed in forecasting stock market trends. Additionally, this paper encompasses recent advancements in stock market prediction systems and the scholarly literature around these advancements. Sentiment analysis, time series analysis, and statistical methodologies are all encompassed within this particular genre. The authors of

 Received: November 2, 2023 | Accepted: December 2, 2023

this article offer a comprehensive examination of these domains. Every year, a substantial number of articles are released across several subfields within the domain of stock market forecasting. The accumulation of scholarly articles over time is noteworthy. There is a need for survey studies that comprehensively examine prevailing research patterns and prospective avenues in the discipline. The primary aim of this paper is to present a coherent framework for researchers and investors to build a robust methodology for forecasting stock prices.

The paper's structure: Section 2 includes research study selection materials and methodologies. Section 3 describes the existing stock price prediction model-building methods used by several academics. Section 4 surveys prior studies that presented multiple share price prediction models. The majority of studies use AI-based models. This study's research design and analysis are discussed in Section 5. The article concludes in Sect. 6.

1.1. Selection of Articles

This investigation included AI-based share price prediction publications. The authors selected articles using keywords like "share price," "predict," "forecast," "sentiment analysis," "financial news," "time series," "economics," "artificial intelligence," "machine learning," and "deep learning.". Research studies published after 2011 were only evaluated, and after that, duplicate publications were excluded, followed by the elimination of studies based on non-suitable titles and abstracts. After these two procedures, research studies can be assigned full-text to decide their inclusion or exclusion.

1.2. Contribution of the Study

Significant study contributions are included below:

- The report reviews artificial intelligence-based share price prediction research.
- The study also tests different deep transfer learning share price prediction methods.
- While all models accurately predicted share prices, the Decision Tree and Stacked LSTM models outperformed others, achieving an average of 93.2% in news classification and 94.7% in price direction and accuracy, respectively.
- The present study uses AI to improve investor decision-making.

2. Materials and Methods

This study followed PRISMA recommendations. The autrhors searched the full-text archives of four publication databases (ScienceDirect, Scopus, IEEE Xplore, and Web of Science) for English-language articles published between 2011 and 2023. Using “stock predict” or “share predict,” “sentiment analysis,” “forecast,” “time series,” “artificial intelligence,” “machine learning,” and “deep learning” and combinations thereof. It returned 951 journal and conference articles. Studies that match Table 1 criteria are accepted or refused.

Table 1 Inclusion and exclusion criteria.

S. No.	Parameter	Inclusion Criteria	Exclusion Criteria
1	Period	Between 2011 and 2023, research studies were carried out.	Articles published before 2011
2	Investigations	Research studies are focusing on the benchmark data-set, proposed methodologies, and the research’s research results.	Research studies focusing on other than share market
3	Compactor	Research studies aimed stock market prediction	Research studied aimed (a)Other prediction model except stock market pre- diction (b) Outcome not of interest (c) Population not of interest
4	Methodology	Research studies focusing on ML or DL techniques	Research studies are focusing other than ML-DL techniques
5	Design of study	Original articles comprised of experiment results	Survey papers, case studies, non-English

As indicated in Figure 1, the PRISMA flowchart can be used at several points in selecting research papers.

After eliminating duplicate publications, only research articles published between 2011 and 2023 are considered in the second screening. The authors only researched after 2011 because AI is widely utilized to predict stock direction and price movement. Later rounds of research paper production removed abstracts, titles, and readings. This is when full-text readings are done, and studies are accepted or rejected depending on suitability. The systematic review process concludes with selection when selected research papers are considered for inclusion based on all inclusion criteria. Each of the listed studies included the results of an experiment. A literature survey provides a concise overview of these scholarly articles.

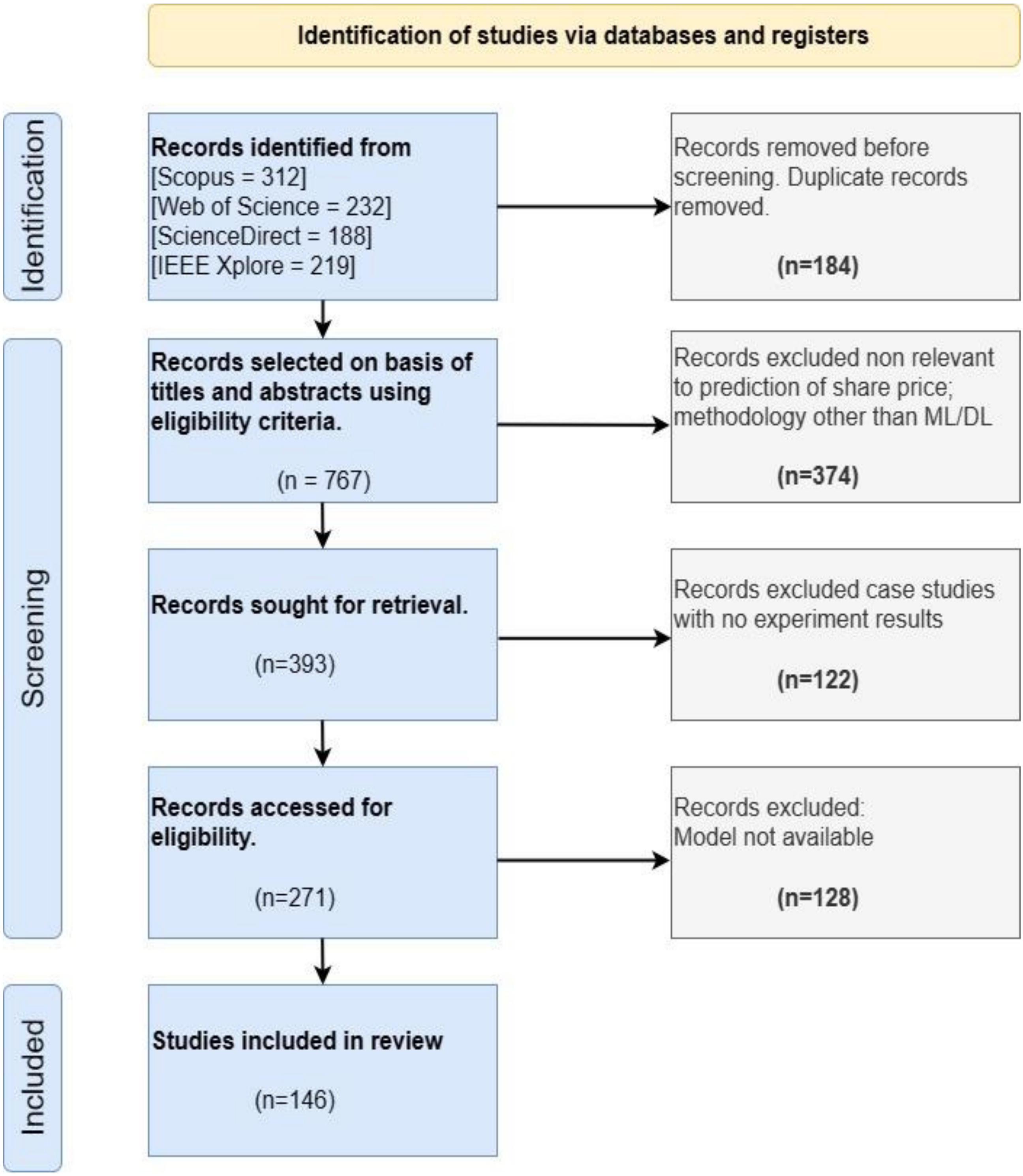

Figure 1 Flow Chart of Search Process.

2.1. Quality Assessment

This study specified certain quality evaluation constraints for research publications. Inclusion/exclusion criteria affect study relevance. All papers in this review use machine or deep-learning stock market prediction systems.

2.2. Investigations

The study conducted various investigations, which are outlined as follows:

• **Investigation 1:** Which ML and DL methods are widely preferred to anticipate share price direction, movement, and price?

• **Investigation 2:** Which share market data sites have been extensively studied, and which year have most share price prediction research been published?

• **Investigation 3:** Which criteria evaluate ML and DL models?

3. Existing framework for Stock Market Prediction

Stock markets have been studied for years to identify trends and predict changes. Two key stock market forecasting hypotheses are the EMH (Yalccin, 2010) & RWT (Oskooe, 2011).

3.1. Efficient Market Hypothesis (EMH)

"Market Efficiency" means the financial literature's correlation between news and stock price. Fama defines "Market Efficiency" in his 1998 work (Fama, 1998), dividing it into three groups: weak, medium, and strong. In Figure 2, the rings reflect each EMH structure's information volume. Note that the weak structure covers minimal data, while the strong form covers all data. Additionally, every advancing structure integrates the previous ones.

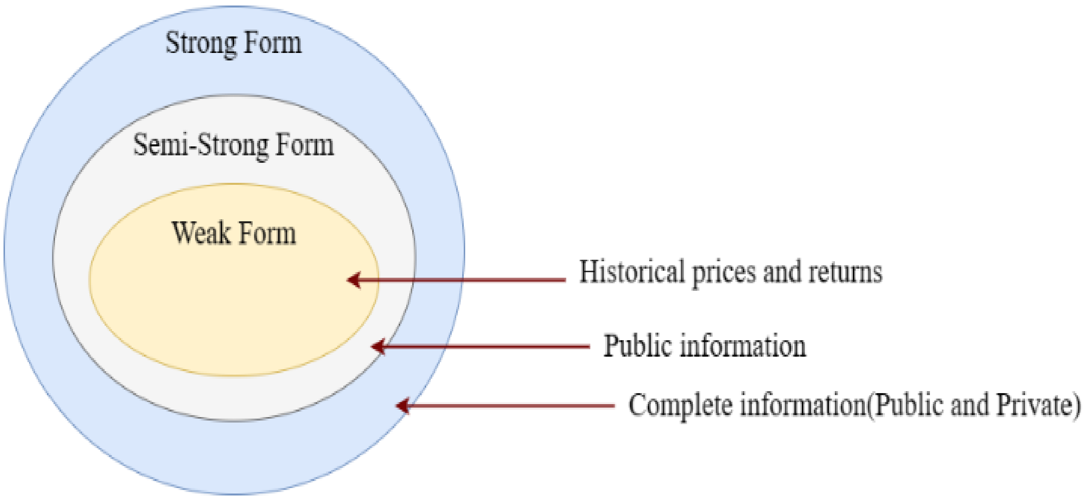

Figure 2 Representation of Efficient Market Hypothesis.

The EMH suggests news drives stock prices, not fundamentals. Continuous information makes stock price predictions impossible with more than 50% accuracy.

• **Weak-form of efficiency:** This implies that prior stock rate data is already factored into the current stock price. Technical analysis did not determine trading decisions for investors to earn from the market. Fundamental analysis aids in identifying cheap or overvalued stocks for trading decisions with higher returns compared to the market.

• **Semi-strong form of efficiency:** In this, past share price record & public record are already factored into current stock prices. Semi-strong form efficiency suggests that only investors with private information may profit from market profits, as neither technical nor fundamental analysis can do so.

• **Strong form of efficiency:** Strong efficiency leads to no additional earnings and wastes time in the market. This combines weak & semi-strong forms with public & private data and historical pricing history. This implies that all available data is priced at the market price. The EMH may need to be more accurate in a workout as many market participants consistently outperform the market. The validity of EMH is disputed as there is ample evidence that markets are not entirely efficient (Gupta et al., 2008). Predicting future stock prices or indices is more possible than random trading (Shahvaroughi et al 2021). Market participants prioritize EMH in forecasting due to its inconsistency with other methods. If EMH is true, random guessing outperforms technical foundations and time series analysis in predicting outcomes. Stock market disasters such as Oct. 1987, Oct. 2002, Oct. 2007, Sept. 2008, June 2015, and Mar. 2020 contradict the EMH as they may not depend primarily on random information.

3.2. Random Walk Hypothesis (RWH)

Vaga (1990) introduced the random walk concept. The writer claimed that value variations among periods are meaningless. Its deviation also depends on the time interval between frames. According to the random walk hypothesis (RWH), current market prices are the most reliable forecasters of future prices, with a stochastic error term (Pant et al., 2001). Past and present stock prices are uncorrelated, according to Random Walk theory. In the study, the authors investigate the random walk hypothesis in the Indian stock market. The heteroscedasticity- adjusted variance ratio test rejected the random walk hypothesis for Indian daily and weekly market indices. The study revealed that indices change daily, causing time-varying volatility and excluding rare trades, rejecting the RWH. Due to stock price mean reversion, the random walk null hypothesis is rejected. The study confirms the mean reversion and overreaction of Indian stock prices. This allows traders/investors to predict prices and profit. A study (Lee, 1992) found strong evidence of RWH rejection, highlighting the need for market forecasting studies.

3.3. Fundamental Analysis

The fundamental analysis follows RWH and EMH as a second-category concept. This school of thought originated from John B. Williams' firm-foundation theory, which used the discount cash flow (DCF) technique (Williams, 1938). Later, works like Security Analysis (Hoffman, 1935) popularized long-term investing education. Fundamental analysis considers financial status, board of directors, annual report, balance sheets, profit/loss accounts, topography, climate, and politics to predict future stock prices (Silva et al., 2020; Jing et al., 2021; Agrawal et al., 2022; Singh and Malhotra, 2023). Fundamental analysis must be flexible since factors are unstructured. Some academics believe machine learning has made stock market prediction easier, utilizing unstructured data. Long-term price movement can be predicted using fundamental analysis, but not for short-term fluctuations (Khan et al., 2020). Fundamental analysts use publicly available firm, economy, and industry data to analyze stock price movement in three dimensions. Fundamental analysts also consider business financial ratios. As indicated in Fig. 3, the fundamental analyst ratio is used to understand the share market in three dimensions: Firm analysis, Economy analysis, and Firm analysis.

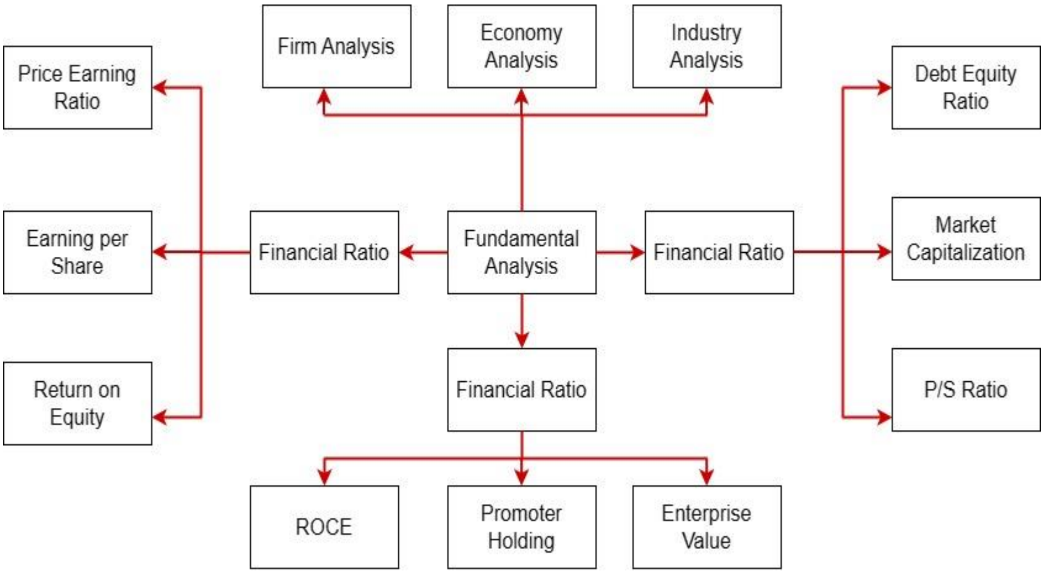

Figure 3 Ratios involve in fundamental analysis of the organization in stock index.

- **Market capitalization (MC):** MC counts all of the shares shared in the market. MC categorized the shares into small, medium, and large. MC can be calculated using the following formula:

$$MC = Total\ Shares * Share\ Price$$

- **Price/Sales ratio (P/S Ratio):** Metric check whether the share price accurately reflects the stock's value. P/S is calculated as follows:

$$\frac{P}{S} = \frac{Price\ of\ Share}{Return\ over\ twelve's\ time\ frame}$$

- **Price/Book ratio (P/B Ratio):** This has to do with determining how much stock is worth compared to its market value. This ratio is a measure of underestimation or overestimation of a stock. P/B is computed as:

$$\frac{P}{B} = \frac{Share\ Price}{Whole\ assets - Tanglible\ assets\ and\ liability}$$

- **Earnings per share (EPS):** A company's profitability measure is net income divided by outstanding shares. Following equations show how EPS can be calculated in two different methods.

$$EPS = \frac{Net\ Income}{Outstanding\ Shares}$$

$$Weighted\ EPS = \frac{Income\ after\ tax - Dividends}{Total\ number\ of\ remaining\ stocks}$$

- **Return on assets (ROA):** It means the amount of money a company makes about the total amount of money it has. The metric that quantifies a corporation's profitability about its total assets or resources.

$$ROA = \frac{Net\ Income}{Total\ assets}$$

When interest on debt is not considered, ROA is calculated using the following formula:

$$ROA = \frac{Net\ Income + Interest\ Expenses}{Average\ of\ total\ assets}$$

- **Return on Equity (ROE):** It shows how well the shareholder's money was spent and how much profit it generated. A low return on equity indicates that the shareholder's money was not adequately invested. ROE is calculated using the following formula:

$$ROE = \frac{PTP}{SE}$$

- **Debt/Equity Ratio(D/E):** This shows the relative impact of available and used capital. A low debt-to-equity ratio indicates that all available credit was not used. Next, the authors explain how to calculate the debt-to-equity ratio (D/E).

$$\frac{D}{E} = \frac{Overall\ Liabilities}{Shareholder'sEquity}$$

Recently, text mining tools have made stock market analysis more popular. Numerous studies have employed fundamental analysis for stock prediction. However, Kim et al. (2020) highlights weaknesses in text mining that impact decision- making effectiveness. Multilingual text minor change dependency is a significant challenge with limited solutions supporting multiple languages (Zhang et al., 2017). Scholars Sharaf et al. (2021); Kumar et al. (2021); Wang et al. (2018) maintain that the stock market remains unpredictable despite technical and fundamental analysis predictions.

3.4. Technical Analysis

The objective of the technical analyst is to forecast fluctuations in the stock market by employing price charts and technical indicators (Picasso et al., 2019; Zhang et al., 2019; Chandar, 2019; Pant and Bishnoi, 2001; Idrees et al., 2019; Shukla and Singh, 2019). Figure 4 shows how the predictive algorithm computes indicators from historical stock values, which are pre-owned. The technical analysis encompasses the examination and utilization of several technical indicators mentioned in studies (Haq et al., 2021; Ji et al., 2023) as depicted in Figure 4.

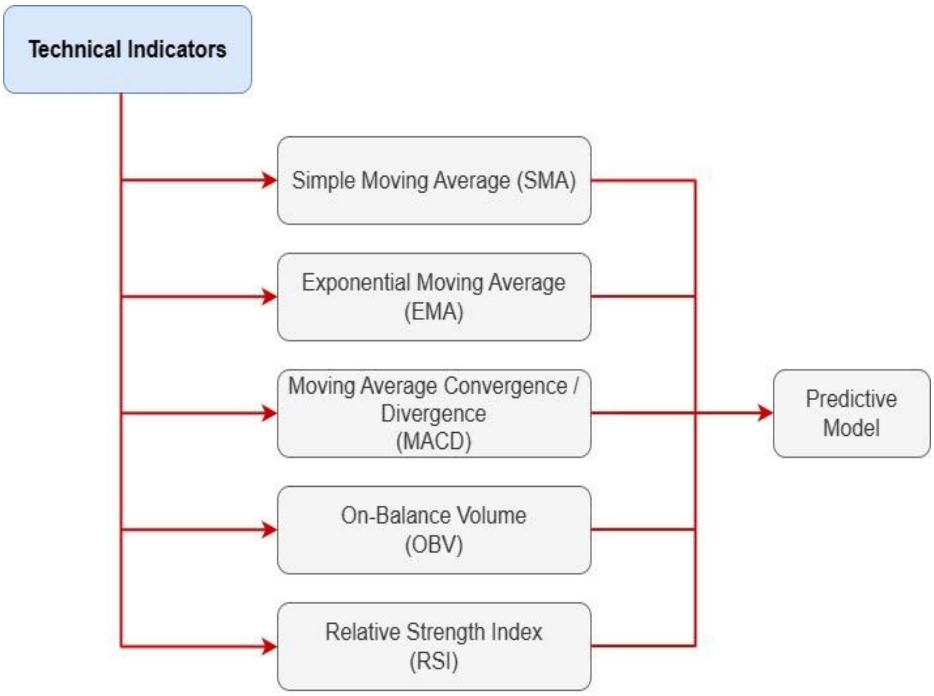

Figure 4 Technical Indicators.

• **SMA:** The simple moving average (SMA) is determined by the summation of the closing prices of a stock over a specific number of periods, denoted as “n,” and then dividing this total by the number of periods to calculate the average value of the SMA (Dinesh et al., 2021).

• **EMA:** However, the current day’s EMA computations depend on the previous day’s SMA calculations.

$$EMA = Price(t) * k + EMA(y) * (1 - k)$$

In the above equation, parameter t represents today’s EMA, whereas entity y represents yesterday’s EMA. Despite this, N represents the total number of days in EMA followed by k represents smoothing which is equal to 2/ (N + 1).

• **MACD:** Short-term and long-term patterns are compared in the MACD indicator to make predictions about stock market developments. MACD can be calculated by comparing a 26-day EMA (EMA) with a 12-day EMA.

$$MACD = \sum_{i=1}^{n} EMA_k - \sum_{i=1}^{n} EMA_d$$

that is, k = 12 is the number of EMA days and d = 26 reflects the number of days in EMA.

• **OBV:** It utilizes the analysis of volume-flow to predict fluctuations in stock market prices. A decreasing On-Balance Volume (OBV) line signifies a downward movement in the stock price, while an increasing OBV line suggests a potential future rise in the stock price.

• **RSI:** Indicator that measures if a stock has been oversold or has been overbought. Is shown in the following equation.

$$RSI = 100 - \frac{100}{1} + RS$$

RS represents the average gain/average loss

3.5. Behavioral Analysis

Behavioral finance is a branch of economics that studies how investors and financial professional’s cognitive biases and mental models influence their decision-making processes. Psychological variables and biases may underlie market anomalies, particularly in the stock market, manifesting as unanticipated price fluctuations. The reference provided is (Göçken et al., 2016) on herding. Due to its significance in the investing sector, the Securities and Exchange Commission (SEC) employs professionals specializing in behavioral finance.

The study of behavioral finance encompasses a diverse range of perspectives for analysis. Although other considerations exist, the stock market is regarded as a domain within finance wherein psychological behaviors are believed to influence market outcomes and returns. The classification of behavioral finance sheds light on the influence of many factors on market performance (Picasso et al., 2019; Ready, 2002).

Within behavioral finance, it is widely acknowledged that investors do not adhere strictly to rational decision-making processes. Instead, they are believed to be influenced by psychological factors and exhibit tendencies towards self-control within the range of normal human behavior. Individual’s mental and physical health can substantially influence their ability to make prudent financial choices. The attitudes of investors may be subject to alteration based on the improvement or deterioration of their health conditions. Due to this phenomenon, individuals experience a decline in their rationality and efficacy when confronted with money or other practical circumstances.

The examination of biases holds significant importance within the realm of behavioral finance. The origins of bias may encompass a diverse range of factors. Five primary categories of prejudice exist under which most cases might be classified (Waweru et al., 2008; Seetharam, 2013). By recognizing and avoiding different behavioral finance fallacies, an investor can improve their ability to focus on a particular company or industry.

The majority of stock market prediction methodologies utilize technical analysis as a means to forecast patterns in stock values. However, according to a study conducted by (Souma et al., 2019), it is argued that quantitative data may not accurately capture the wide range of financial circumstances experienced by different organizations. In the contemporary era of social media, the inclusion of unstructured data derived from conventional news sources and social net- works can serve as a valuable complement to quantitative data, enhancing the accuracy and efficacy of forecasting methodologies.

- **Mental accounting:** Individuals often exhibit a cognitive inclination to track and analyze their expenditure patterns mentally.
- **Herd behavior:** Based on the theoretical framework of herd behavior, it can be posited that a significant proportion of individuals tend to emulate the financial actions of the collective. Herding activity is frequently cited as contributing to significant fluctuations in the stock market.

- **Emotional gap:** Powerful emotions such as anxiety, fury, fear, and exhilaration have the potential to precipitate impulsive decision-making. The lack of emotional regulation among individuals significantly contributes to their proclivity for illogical decision-making.
- **Anchoring:** The anchoring phenomenon arises when an individual establishes a cognitive connection between a specific level of spending and a well-known reference point. One illustrative instance involves adhering to a predetermined financial plan, while another entail rationalizing expenditures by considering the relative significance of different preferences.
- **Self-attribution:** Self-attribution refers to the inclination to engage in actions based on an inflated assessment of one's skill level or comprehension. The capacity to ascribe personal achievements to one's abilities is commonly associated with inherent aptitude. In this group, people tend to overestimate their competence, especially when contrasted with others.

In the paper (Mahmood et al., 2020), investing decisions and their resulting results are examined to ascertain the influence of behavioral elements. The primary data was obtained by distributing questionnaires to individuals who regularly engage in activities within the Pakistan Stock Exchange. The study also examined the effect of financial literacy as a moderator variable in the relationship between behavioral traits and investment decision-making, as it has been found to assist investors in mitigating behavioral biases and making rational choices that enhance their wealth. The AMOS software analyzes and evaluates data and does hypothesis testing. The study's findings revealed that the investment decisions made by investors were predominantly influenced by behavioral variables such as heuristics, prospects, markets, and herding. The study's findings indicate that prospects and herd mentality positively influence investment success but are negatively impacted by heuristics and market conditions. The findings indicate that the investment decision-making process of investors is more strongly influenced by heuristic, prospect, and market factors when they possess a higher degree of financial literacy. Conversely, the impact of the herding variable on investment choices is less pronounced among investors with lower levels of financial literacy.

The most effective approaches for predicting stock market trends seem to be fundamental and technical analysis in conjunction with sentiment analysis and deep learning models. The outcomes of employing an ensemble technique for forecasting exhibit a significant degree of promise. As documented in academic literature, Figure 5 illustrates the algorithms used in predicting stock market fluctuations.

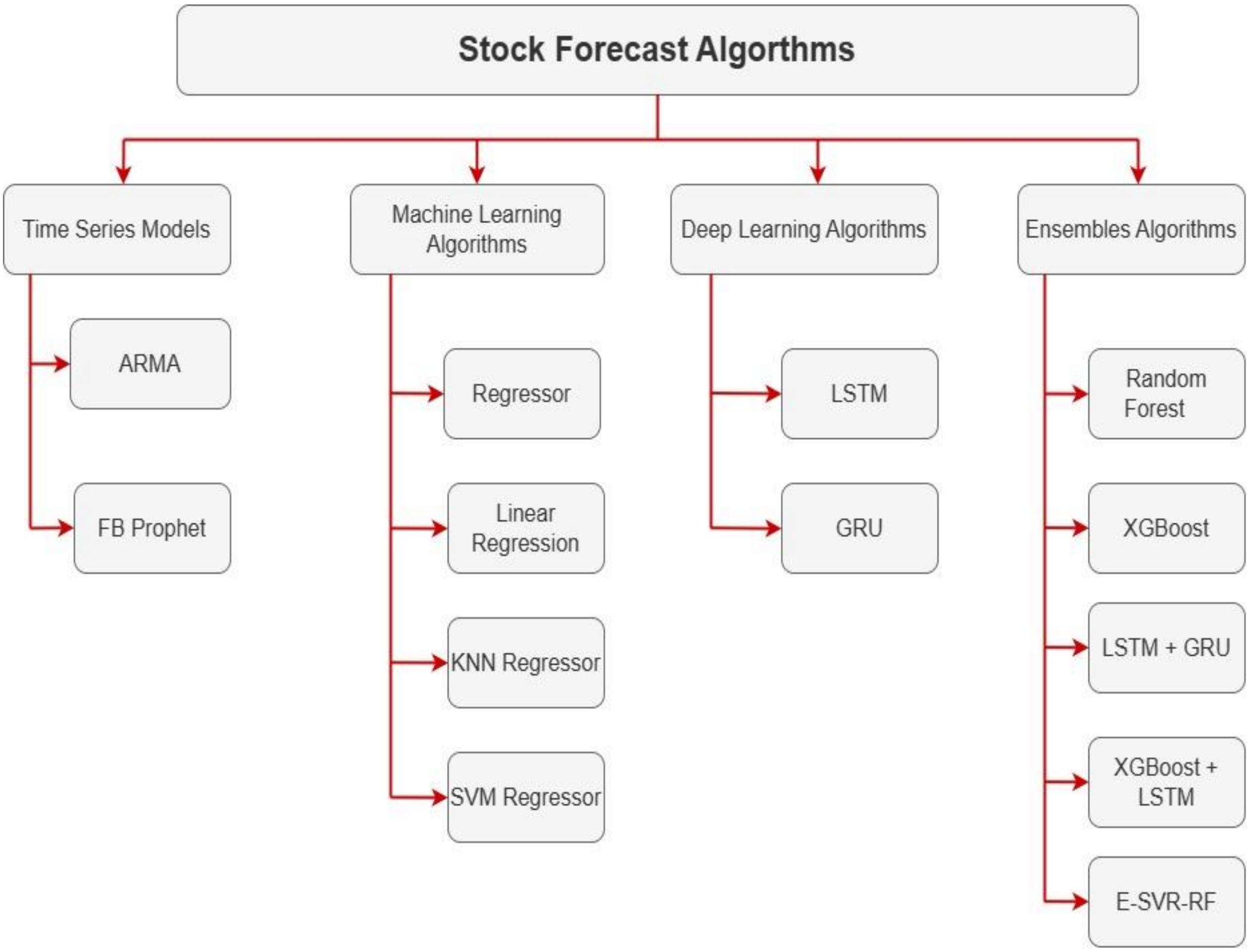

Figure 5 Stock Forecast Algorithms.

3.6. Time Series Algorithms for stock market applications

Stock price data can be considered a representative time series instance. This section covers several well-established methodologies, including the Auto Regressive Moving Average (ARIMA) and the FB Prophet. Before the emergence of deep learning models, these strategies were commonly utilized.

3.6.1. ARIMA

Auto-regressive integrated moving average (ARIMA) is a statistical analysis model for understanding and forecasting time series data. An auto-regressive statistical model makes predictions by analyzing the past. Earnings estimates or stock price forecasts can benefit from using an ARIMA model, which uses historical data to make predictions. One may quickly grasp the following parts of an ARIMA model:

Auto-Regression (AR): "Auto-Regression" pertains to a theoretical framework wherein a variable susceptible to change is regressed onto its previous values.

Integrated (I): To stabilize a time series, the data values are substituted by the difference between the data values and their respective preceding values.

Moving Average (MA): To elucidate the relationship between a residual error derived from a moving average model and a specific observation inside a data set characterized by lagged values.

Each component of the ARIMA model operates as a parameter, adhering to a standard syntax. ARIMA models are commonly denoted as ARIMA (p, d, q), where p, d, and q represent integer values that specify the particular characteristics of the ARIMA model employed. These are the limits:

p: The measure of the number of observations preceding the current one in the model.

d: The frequency at which variations occur in the raw observations, indicating the extent of divergence.

q: The moving average order refers to the window size utilized for calculating the average.

In the context of a linear regression model, the composition of the model is determined by the quantity and nature of the components included. A parameter value of zero (0) signifies that the model should disregard the component under consideration. Hence, an ARIMA model can be formulated to fulfill the same purpose as an ARMA model or a simpler AR, I, or MA model. The authors know that converting the series into a stationary time series is a precondition for applying the various models. A difference or integrated approach is utilized to accomplish this task, whereby the value at time t minus one is subtracted from each of the t values inside the time series. The second-order difference is employed when the initial difference fails to generate a stationary time series. The ARIMA model can be seen as an extension of the ARMA model, incorporating an additional component represented by the letter "I" to account for differences. In summary, the ARIMA model integrates several factors, including past lags, residual errors, and other adjustments to ensure stationarity to predict future values. Auto-regression (AR) and moving average (MA) are integral constituents of the Auto-Regressive Integrated Moving Average (ARIMA) model. is shown in Equation (13),

$$y^i(t) = k + \beta_p * wD_{y_{t-1}} + \cdots + \beta_p * wDy^i_{t-1} + \theta_1 * \epsilon_{t-1} + \cdots + \theta_q * \epsilon_{t_{i-q}} + \in_t i$$

where p represents the auto-regressive model's given degree, D is the degree of different orders, and q is the moving average's given degree.

3.6.2. FB Prophet

Facebook Prophet is a proprietary library developed by the corporation for time series forecasting. When encountering null values in the data set, using FB Prophet can enhance the results' reliability. The formula utilized in FB Prophet has been disclosed in equation:

$$yt = l(t) + sp(t) + \vartheta(t) + \epsilon t, yt = g(t) + s(t) + h(t) + \epsilon n$$

where l(t) represents the linear trend, sp(t) represents seasonal patterns, v(t) represents holiday effects, and ∈n is the white noise error.

3.7. Machine Learning Algorithms for stock market applications

3.7.1. Linear Regression

Predicting a stock's future price by linear regression necessitates constructing a model that incorporates one or more variables, such as the stock's historical closing, opening, and volume prices, among other relevant factors. The primary objective of regression analysis is to model the linear relationship between a set of independent variables and a dependent variable. The linear regression model calculates a best-fit line describing the connection between the independent and dependent variables.

$$O = S_x + k$$

where O is the output, S_x represents the slope, and K is a constant in this approach, Equation generates a linear representation that effectively intersects with a maximum number of data points. A mathematical procedure is employed to fit a straight line between these points to decrease the square of the distance or difference between each point on a graph representing the values in a data set. The prediction of the value of y at a specific x can be made using the hypothesis line. The concept of "linear regression" refers to a forecasting technique. The evaluation of results is conducted through the utilization of many metrics, including the Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Squared Error (MSE), and the correlation coefficient (R-squared) (Gururaj et al., 2019; Dospinescu and Dospinescu, 2019).

3.7.2. K-Nearest Neighbor (KNN)

The K-nearest Neighbors (KNN) algorithm, a classification and regression approach, has been called the "lazy learner" due to its characteristic of not requiring an extensive learning period. The K-nearest Neighbors (KNN) algorithm offers several advantages due to its comparatively straightforward implementation compared to other machine learning algorithms. The K nearest Neighbors (KNN) algorithm does not require additional user input other than specifying the value of K and selecting the Euclidean distance as the distance metric. The aforementioned algorithm demonstrates superior efficiency compared to alternative algorithms despite its relatively slower learning rate due to its heightened stability. The approach may exhibit limited generalized to large-scale data sets due to its omission of the learning step (Tan, 2008). This approach calculates the Euclidean distance. In Equation (16)

$$D(h_i, p_r) = \sqrt{\sum_{i=1}^{n} (p_r - h_i)^2}$$

where P_r represents the predicted value and h_i

3.7.3. Support Vector Machine

A support vector machine (SVM) for stock price prediction can be deemed very efficacious due to its dual functionality as a classification and regression technique. When comparing Support Vector Machines (SVM) to other approaches such as "Peeling + SVM" and "CC + SVM," it becomes evident that employing advanced SVM techniques can enhance the accuracy of predictions (Grigoryan, 2017). Utilizing a delimiter, a support vector machine uses a supervised learning approach to categorize data effectively. The division becomes apparent in the initial processing phases when mapping the data to a feature space with many dimensions. The process involves classifying data points inside n dimensions to identify the hyperplane that maximizes optimality. The data points are categorized into clusters based on their proximity to the hyper-planes. (Bustos, 2017) have demonstrated that the efficacy of the Support Vector Machine (SVM) algorithm may be improved by meticulously adjusting its parameters, including those related to regularization, gamma, and kernel. Likewise, Support Vector Machines (SVM) can assess investor mood, potentially leading to consequential impacts on the market. The model has strong performance on data sets of varied sizes, regardless of their differing dimensional levels.

3.8. Deep Learning for stock market applications

Deep learning models are widely utilized in numerous scientific and technology domains. Artificial neural networks are extensively utilized in stock price forecasting and trend prediction because they can capture intricate patterns, handle substantial volumes of data, acquire and depict features, and adapt to changing market dynamics. This article examines frequently utilized deep learning models in banking and finance. According to (Goodfellow et al., 2016), deep learning is a machine learning methodology that relies on interconnected networks of basic units, allowing computers to acquire intricate concepts using multiple layers of graphically connected nodes. Deep Belief Networks (DBNs) are a notable instance of probabilistic or Bayesian models that have played a significant role in the resurgence of Deep Learning (DL). DBNs consist of nodes representing random variables that exhibit probabilistic associations with each other. Artificial Neural Networks (ANNs) have gained significant popularity in recent years, primarily due to their utilization of nodes that function as neurons. All of the architectures examined in this survey employ Artificial Neural Networks (ANNs), and this section presents an elaborate account of each of these structures. Artificial Neural Networks (ANNs) are computational models inspired by the structure and functioning of the human nervous system and brain, utilized for information processing and problem-solving. Due to their composition of numerous elementary processing components possessing adaptive capabilities, these systems can handle substantial quantities of data concurrently. Given that neurons serve as the brain's essential units for information processing, Artificial Neural Networks (ANNs) are constructed around a significantly abstracted representation of a neuron. The similarities between Artificial Neural Networks (ANNs) and the human nervous system have been extensively discussed in the literature (Jing et al., 2021). These similarities can be summarized as follows:

• Neurons, nodes, and basic computational units are foundational components of all information processing systems.

• Neurons possess the ability to engage in communication both among themselves and with their external environment.

• Neural networks are formed when neurons interact with one another.

• Synapses serve as the anatomical structures facilitating the transmission of information between neurons.

• The amount of information stored inside a neural network is directly correlated with the synaptic efficiency, which may be measured by the weight values or strengths of the synapses.

• During learning, the aggregated weight values, also known as connective strengths, undergo modifications in response to external stimuli.

Deep Reinforcement Learning (DRL) is a deep representation of Reinforcement Learning (RL), categorized as model-based, model-free, or a combination of both. The stock market can possess a Dynamic Recursive Lag (DRL) trait, as it demonstrates the ability to record past states and events through present states effectively, and future states can be projected based on the current state. Consequently, Deep Reinforcement Learning (DRL) has gained significant traction as a prominent methodology for advanced quantitative research within the domain of stock market analysis. In many scenarios, Deep Reinforcement Learning (DRL) can be applied for a range of objectives, such as selecting lucrative stocks or strategically allocating portfolios, as well as devising an optimal liquidation strategy through market simulation (Bao and Yang, 2008; Wang et al., 2018).

3.8.1. LSTM

Recurrent Neural Networks (RNNs) are a type of deep learning algorithm, and LSTM is an improved version of RNNs. Because it can retain the data sequence, an LSTM model can process long sequences of data units and use that information as input in the future. A typical LSTM neuron is depicted schematically in Figure 6. The input gate, the forget gate, and the output gate are its three constituent parts. Each of the gates uses the sigmoid activation function. Following equations provide a mathematical representation of all of the LSTM's gates.

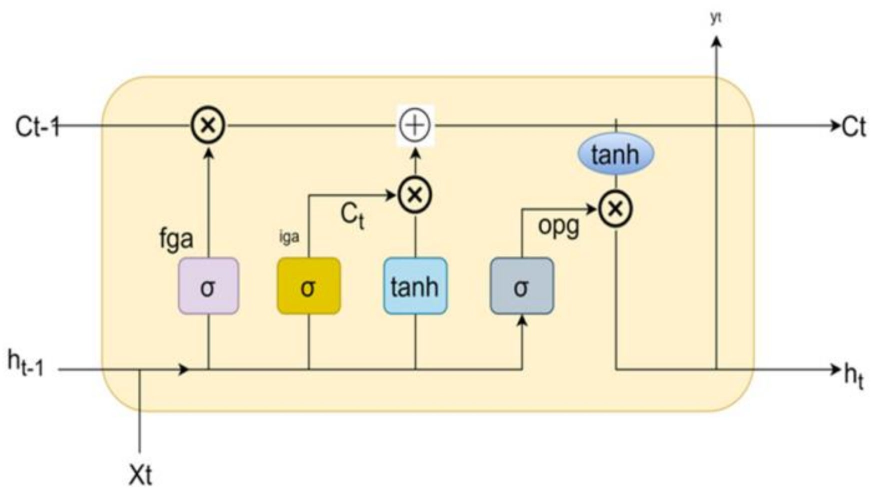

Figure 6 Structure of LSTM model.

- Input gate

$$i_g a = \sigma(W_i p[h_{t-1}, X_c] + b_i)$$

- Forget gate

$$f_g a = \sigma(W_i g[h_{t-1}, X_c] + b_f)$$

- Output gate

$$o_g a = \sigma(W_o p[h_{t-1}, X_c] + b_o)$$

where σ is sigmoid, W_x is the neuron gate (x) weight, h_{t-1} is the result of the preceding LSTM block, X_t is the input, and b_x is bias. Fig. 6 demonstrates the model's capability to connect the upper segment of each cell's memory and transport lines, facilitating seamless data transmission between various storage sites. The collection of cells at each LSTM node stores the data stream named(Pramod and PM, 2020). One of the benefits of utilizing LSTM in a recursive network is its ability to facilitate the training process and establish a longrange causal connection by effectively reducing the error rate (Mukherjee and Joshi 2014). In various domains, such as predicting, neural and deep neural networks have demonstrated superior performance compared to conventional machine learning models. Nevertheless, previous research (Gaba et al., 2023) has demonstrated that logistic regression models outperform neural networks in predicting financial hardship.

3.8.2. Gated Recurrent Unit (GRU)

The Gated Recurrent Unit (GRU) is an additional Recurrent Neural Network (RNN) model that shares certain similarities with the Long Short-Term Memory (LSTM) model. The model effectively captures extended temporal relationships among sequential data and exhibits a shorter training time than LSTM. The GRU utilizes gating mechanisms to manage data flow between current and previous time steps. In contrast to the Long Short-Term Memory (LSTM) model, which incorporates three gating mechanisms, the current model employs only two gating tools: a reset gate and an update gate. Figure 7 illustrates the fundamental structure of a Gated Recurrent Unit (GRU).

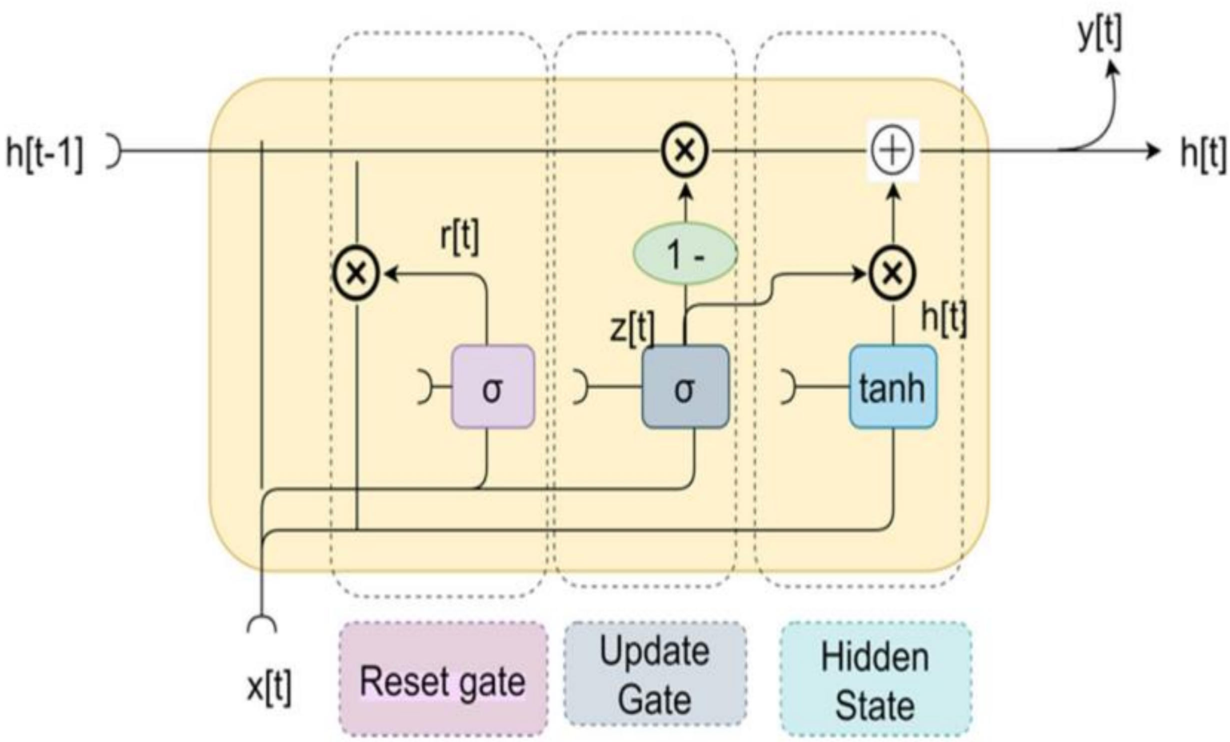

Figure 7 Structure of GRU model.

- Update gate

$$Z[t] = \sigma(W^z x_t + U^z h_{t-1})$$

- Reset gate

$$r[t] = \sigma(W^r x_t + U^r h_{t-1})$$

Where $Z[t]$ is update gate, $r[t]$ is reset gate, σ represents sigmoid function, W_x is neuron gate, U_x is previous weight, h_{t-1} is the result of the preceding GRU block, and x_t is the current input. The Gated Recurrent Unit (GRU) model receives a sequential input of historical stock prices and generates predictions of future stock prices in the same sequential style. During each time step, the input phrase is employed to update the internal state of the GRU network. Subsequently, the final state is utilized to construct the output, as mentioned in the study by (Di Persio and Honchar, 2017) on recurrent networks. LSTM and GRU demonstrate their efficacy in addressing the vanishing gradient problem and integrating deep ensemble approaches.

3.9. Ensembles Algorithms for stock market applications

3.9.1. Random Forest Algorithm

The random forest algorithm employs an ensemble learning approach and falls under supervised learning techniques. This technique demonstrates utility across various scenarios, encompassing classification and regression tasks. The algorithm is founded on the concept of a decision tree (Kaczmarek and Perez, 2022) since it creates many decision trees to generate its output. Figure 8 illustrates the general stages of an algorithm designed for predicting stock market trends and the process of random forest.

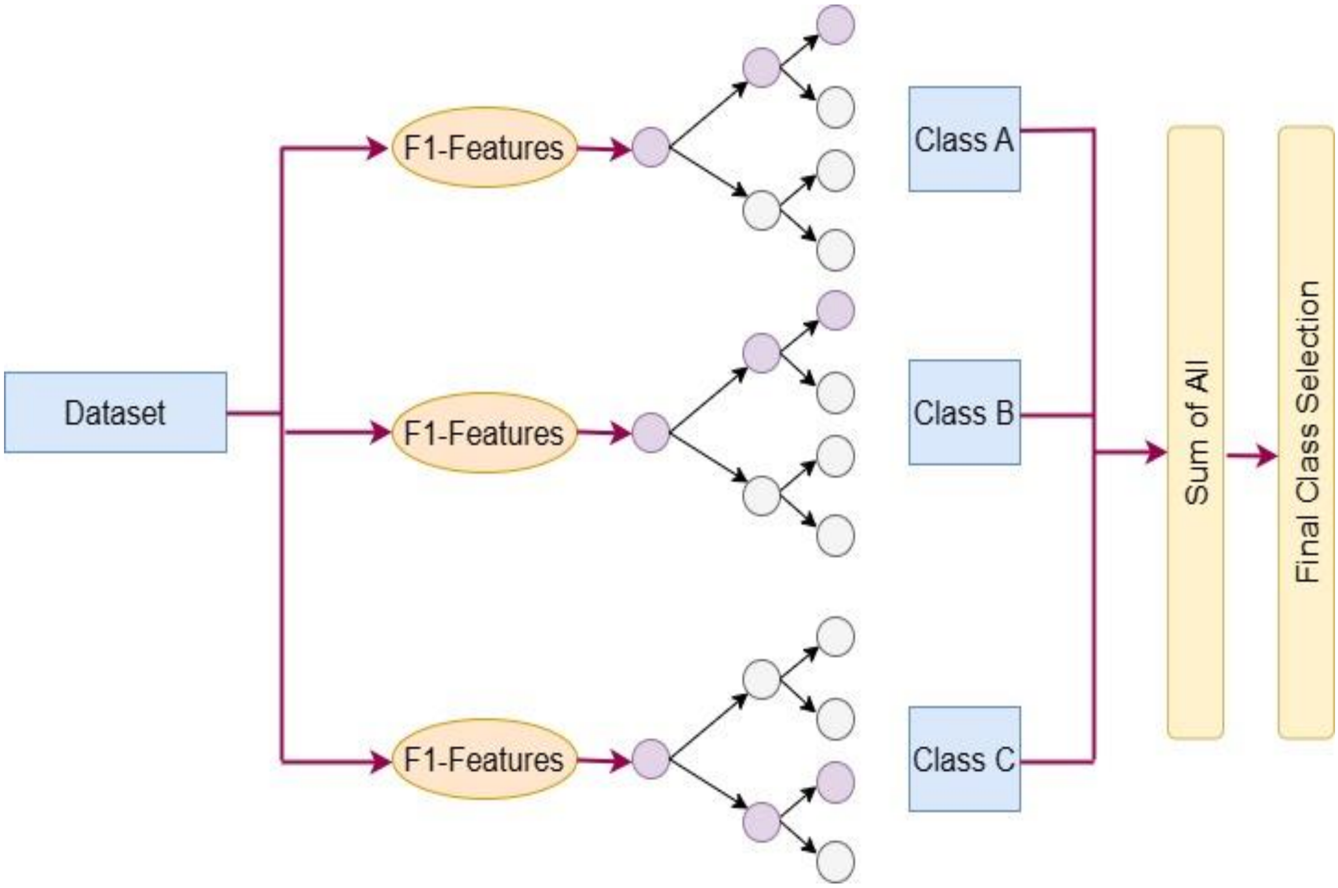

Figure 8 Structure of Random Forest algorithm.

- Step 1: N random records are picke
- Step 2: A decision tree is built based on N inputs.
- Step 3: The number of trees to be considered is picked.
- Step 4: Based on the steps performed before, the output is predicted for each tree.

Nevertheless, the algorithm's computational efficiency may be hindered when dealing with extensive data sets due to the possibility of a significant proliferation of trees (Salles et al., 2018). The random forest method has many potential applications, including predicting stock movements (Souma et al., 2019). The algorithm's accuracy, precision, recall, and F-Score are evaluated compared to SVM, KNN, and logistic regression using assessment parameters derived from the Zagreb Stock Exchange (Kamara, 2022).

3.9.2. XG-Boost Regression Algorithm

XG-Boost is an ensemble method utilized in machine learning, exhibiting similarities to the random forest methodology while still possessing notable differences. The amalgamation of many straightforward learning techniques, such as decision trees, is evident in this approach. The decision tree's weights are iterative adjusted using the gradient, making it a valuable prediction model for stock market forecasting (Kumar et al., 2021). Figure 9 illustrates the implementation of the XG-Boost algorithm.

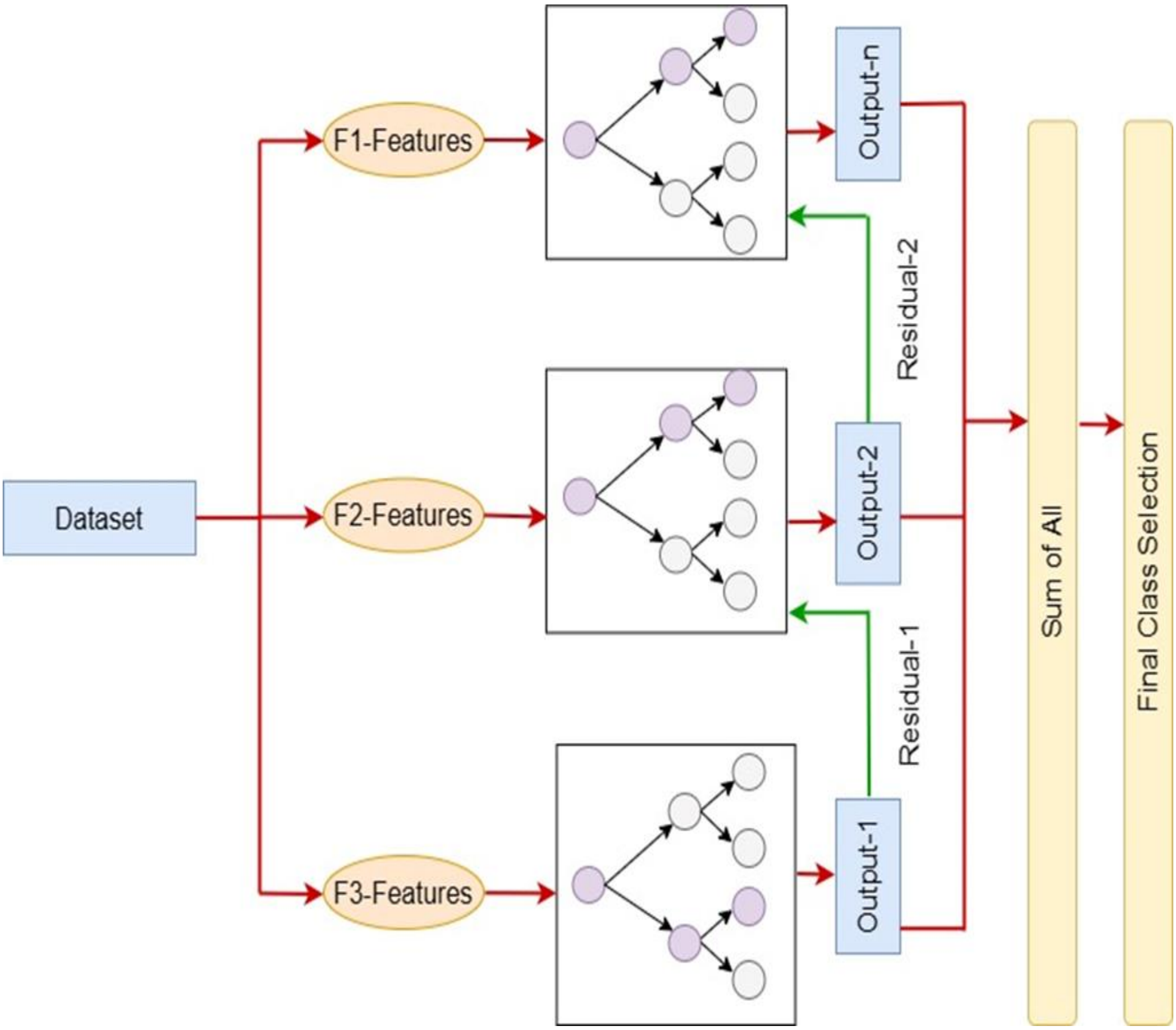

Figure 9 Structure of XGBoost algorithm.

3.9.3. E-SVR-RF (Ensemble Support Vector Machine-Random Forest)

Random forest ensemble support vector regression is an example of an ensemble methodology that employs the bagging strategy. To address the complex interaction across many data groups, the researchers utilized an ensemble learning methodology that consists of a support vector regressor and a random forest. This strategy enables us to provide predictions for each group The bagging technique is subsequently employed to amalgamate various outcomes and present a unified prediction derived from a hybrid model that utilizes a weighted average approach (Xu and Keselj, 2019). The support vector machine employs a hyperplane that maximizes the margin between the predicted and observed prices, enabling it to capture intricate correlations. On the other hand, the random forest technique mitigates the issue of over-fitting by amalgamating decision trees. The academic community has put forth various methodologies to forecast the values of financial instruments. These include the employment of XG-Boost + LSTM, blending ensemble (LSTM + GRU) as suggested by (Li et al., 2021), and the utilization of ensemble techniques for sentiment analysis as given by (Gite et al., 2023). In conclusion, our analysis demonstrates that the increasing use of ensemble methodologies that harness the capabilities of deep learning and machine learning is substantiated.

4. Literature Review

The profitability of technical analysis-based trading methods has been varied in academic studies. Some have shown that technical tools perform poorly, as suggested by (Di Persio and Honchar 2017; Singh 2023; Singh and Shukla 2021; Vargas et al., 2018) while (Sullivan et al 1999; Oncharoen and Vateekul 2018; Wang et al 2023) enthusiastically encourage them to provide higher returns.

Technical analysis analyses price changes with various instruments. Many academics have studied new technical tools, but only some have concentrated on traditional tools like candlesticks and chart patterns. This study focuses on candlestick charting patterns, head-and-shoulders pricing patterns, and classic technical analysis methods. The study examines the performance of stock market indices in India, considering academic critique of technical analysis and the importance of passive investing. To achieve these objectives, the literature review is divided into four phases.

- Studies analyzing the use and profitability of technical analysis.
- Studies analyzing the performance of statistical approaches on various stock market indices globally.
- Studies analyzing the performance of sentiment analysis of stock news across the globe.
- Studies analyzing the performance of AI approaches on various stock market indices across the globe.

4.1. Studies Analyzing the Use and Profitability of Technical Analysis

In the paper (Yazdi et al., 2013), the authors examined the effectiveness of the forex market's MACD indicator. They examined the Euro, British Pound, Franc, and Japanese Yen hourly data from January 2001 to December 2010. Despite being a commonly utilized technical indicator, MACD only worked for the Euro and lost for the British Pound, France, and Japanese Yen.

In the paper (Wang et al., 2015), the researchers conducted an empirical analysis on a data-set of 7,846 Chinese technical trading rules. The results indicated that these criteria exhibited superior performance in volatile markets while demonstrating inferior performance in stable markets. Their ability to forecast the 2005-2007 financial bubble was superior.

In the paper (Patel et al., 2015), the authors examine four prediction models, namely Artificial Neural Network (ANN), Support Vector Machine (SVM), Random Forest, and Naive-Bayes. The evaluation is conducted using historical data spanning from 2003 to 2012, focusing on the performance of these models in predicting the behavior of CNX Nifty, S&P BSE Sensex, Infosys Ltd., and Reliance Industries within the Indian stock markets. Traders employ a variety of technical indicators, each possessing distinct attributes, to forecast stock movement. The classification models, including Support Vector Machines (SVM), Random Forest, and Naive Bayes, exhibit a high level of accuracy, approaching over 90%.

In the paper (Wei, 2016), the suggested model is evaluated using TAIEX data-sets spanning from 2000 to 2006, as well as Hang Seng Index (HSI) data-sets covering the period from 2000 to 2004. Every year represents a distinct subset of data. The EMD-ANFIS model had superior performance compared to the other models when applied to the TAIEX and HSI data-sets since it yielded the lowest forecasting errors.

In the paper (Zhou et al., 2016), the authors use past prices to predict stock or market index price movements one step ahead. The analysis used 13 technical indicators. Because Chinese stock exchanges have 242-minute trading days, the authors compare the suggested technique against ARIMA, ANN, and GAN-FD with T set to 242.

In the paper (Wu et al., 2021), the study performs predictive experiments utilizing time windows of 1-day, 3-day, and 7-day durations and examines the progression of anticipated results through time. This study examines and contrasts historical, option, and future data with Support Vector Machines (SVM), CNNpred, CNN-corr, and Neural Networks (NN) performance. The neural net- work framework incorporating convolutions and long–short-term memory units demonstrated superior performance in prediction tests compared to statistical approaches, conventional convolutional neural networks, and long–short-term memory models.

In the paper (Xiao, et al., 2020), the communication stock trading program selected Auto-desk indicators daily from a total of 102 trading days in the year 2015 to obtain inventory data. Please include the highest price, lowest price, starting price, closing price, trading volume, 5-day, 10-day, 30-day, and 60-day moving aver- ages for today. The experimental findings demonstrate that ARI-MA-LSSVM exhibits exceptional predictive capabilities, displaying high levels of accuracy and little error across several data-sets, including motherboard, small, medium, and GEM data.

In the paper (Yildirim et al., 2021), the framework was developed to identify the optimal threshold value, denoted as T, that maximizes entropy. This is achieved by constraining the iteration number to the upper bound of the threshold obtained from histogram analysis. The researchers employed two Long Short-Term Memory (LSTM) models to forecast the directional movement of the EUR/USD currency pair. The researchers observed that the baseline models had a higher volume of transactions in their five-day-ahead prediction experiments, however, the hybrid model demonstrated superior accuracy with an average prediction rate of 35.12

In the paper (Chandar, 2022), the incorporation of technical indicators and convolutional neural networks (TI-CNN) played a crucial role in the formulation of a stock trading model aimed at forecasting market movements. The analysis of empirical data indicates that the TI-CNN model exhibits enhanced accuracy and F1 score performance in predicting the entry and exit points of stock prices com- pared to previous deep learning models. When examining the TI-CNN model in comparison to previous deep learning models, a notable distinction can be observed. The assertion is supported by the evidence of the model's higher accuracy and F1 score performance.

The summary of research studies based on technical indicators is shown in Table 2.

Table 2 Summary of studies in stock price prediction using technical indicators.

Reference	Published Year	Data Source	Market Index	Time Frame	Indicators Used
(Wang, 2023)	2023	Stock	China stock market	Daily	SOP, SCP, MACD, EMA, SMA, RSI
(Chandar, 2022)	2022	Stock	Individual stocks	Daily	SMA, EMA, MACD, RSI, ROC, %K, %D, %R
(yildirim, 2021)	2021	Stock	FOREX	1-Day, 3-Days, 5-Days	OHLCV, MA, MACD, ROC, RSI, BB, CCI
(xiao, 2020)	2020	Stock	China stock market	5-Days, 10-Days, 30-Days, 60-days	MA, OHLCV
(wu, 2021)	2019	Stock	US market	1-Day, 3-Days, 7-Days	SMA, RSI, OHLCV
(zhou, 2018)	2018	Stock	China stock market	Daily	SOP, SCP, MACD, EMA, SMA, RSI
(garcia, 2018)	2018	Stock	German Dax-30 index	1-Day ahead	SCP, SOP, MACD, SMA, MTM, OBV, RSI and PSY
(Ibrahim, 2017)	2017	Stock	NSE	Daily	MA, AR
(Chen, 2017)	2017	Stock	SSE	Intraday	SMA, EMA, MACD
(Wei, 2016)	2016	Stock	TAIEX and HIS	Daily	RSI, SMA
(Goccken, 2016)	2016	Stock	BIST100	Intraday	SOP, SCP, RSI, SMA
(Nayak, 2015)	2015	Stock	BSE Sensex and CNX Nifty	1- Day, 1-Week and 1-Month	SMA, RSI, EMA, Triangular Moving average (TMA)
(Patel, 2015)	2015	Stock	CNX Nifty and S&P BSE Sensex	Daily	SMA, RSI, Momen-tum, stochastic, MACD
(Stankovic, 2015)	2015	Stock	S&P	Daily	SMA, RSI, MACDB, EMA
(Bisoi, 2014)	2014	Stock	BSE-SENSEX	1-Day ahead	SMA, SOP, SCP
(Dash, 2014)	2014	Stock	SSE	Daily	SOP, SCP, RSI, SMA
(Fang, 2014)	2014	Stock	SCI	1-Day ahead	EMA, SMA
(Ticknor, 2013)	2013	Stock	Goldman Sachs Group, Inc.	Daily	OHLCV
(Kumar, 2013)	2013	Stock	BSE-SENSEX, NIFTY	Daily	-
(enke, 2013)	2013	Stock	S&P500	Daily	Treasury Bill, SOP, SCP
(Abhishek, 2012)	2012	Stock	Microsoft Corp.	Daily	OHLCV
(shen, 2012)	2012	Stock	NASDAQ, S&P500, DJIA	Daily	-
(Das, 2012)	2012	Stock	NSE	Daily	EMA, SMA
(mohapatra, 2012)	2012	Stock	NSE	Daily	EMA
(argiddi, 2012)	2012	Stock	BSE	Daily	-
(Joseph, 2011)	2011	Stock	US treasury	Daily	US treasury bonds
(Enke, 2011)	2011	Stock	S&P500	Daily	T-Bill rate, Money supply
(khan, 2011)	2011	Stock	ACI Pharmaceutical	8 Days ahead	NAV, P/E, EPS
(Atsalakis, 2011)	2011	Stock	NBG	Daily	SOP, SCP and SMA
(Sureshkumar, 2011)	2011	Stock	NSE	Intraday	SCP, SOP, HL prices
(lu, 2011)	2011	Stock	SSE	Daily	SMA, MV, MVR and EMA

4.2. Studies Analyzing the Performance of Statistical Approaches on Various Stock Market Indices Across the Globe

In the paper (Kumar and Murugan, 2013) the researchers used the SARIMA and XGBoost models to make predictions about the stock market indices. Accuracy is 89.48%, mean absolute error is 15.612%, and mean absolute percentage error is 10.520% in the hybrid model that combines SARIMA and XGBoost. Microsoft stock data is readily available on Yahoo Finance, including the day's low and high prices, the day's opening and closing prices, and the day's adjusted closing price and volume traded. The research presents three new hyper-parameters, AR, I, and MA, that are seasonal parameters of the series. The duration of seasonality is also proposed to be determined by a new parameter.

In the paper (Choy et al., 2021), the present study conducted a comparative analysis of various forecasting models, including ARIMA, LSTM, SARIMA, Holt Winter, and FBProphet. The analysis focused on their performance in predicting stock trends within six distinct Bursa Malaysia stock market sectors. Data on stock prices from sectors like Construction, Technology, and Plantation were gathered from Bursa Malaysia. Among the five-time series models, ARIMA exhibits the lowest average RMSE and MAE, with values of 0.0270 and 0.0189, respectively.

In the paper (Winata et al., 2021), the prediction of Apple stock prices was conducted by researchers employing three SARIMA models, each having respective AIC values of 580.165, 451.591, and 114.612. The optimal model demonstrated a MAPE of 36.05%. The stock data-set of Apple incorporates a collection of 2518 daily stock prices obtained from Kaggle. The data-set exclusively focuses of the "close" attribute to forecast future stock values. The individual in question, referred to as "Dicky", the graph exhibits a state of stationary as the absolute value of the fuller statistic approaches zero. The auto-correlation and partial auto-correlation plots exhibit a significant peak at a latency of 12. The SARIMAX (1,0,1) (2,2,2,12) model produced a MAPE of 36.05% and a RMSE of 112.21%.

In the paper (Qin et al., 2021), the models for the closing prices of JD and PDD stocks from July 26, 2018, through January 29, 2021, were built using the ARIMA model and the BP neural network. The p-value of the Augmented Dickey-Fuller test is statistically significant at a level lower than 0.01. This suggests that the new sequence does not possess a unit root. Based on the analysis, it can be concluded that the ARIMA model with parameters (1, 2, 2) is more suitable for the given time series. The ARIMA forecasting model demonstrates a favorable predictive performance, characterized by a minimal relative error of 0.00231, a statistically significant relative error of 0.03170, a comparatively small relative error of 0.00184, and a relatively substantial relative error of 0.02739.

The summary of research studies based on the statistical approach is shown in Table 3.

Table 3 Summary of studies in stock price prediction using statistical techniques.

Reference	Published Year	Market	Period	Attributes	Models	Performance Metrices	Results
(Kumar, 2022)	2022	Microsoft Stock	2010-2020	OHLCV	SARIMA + XGBoost	MAE, MAPE, Accuracy	MAE: 15.612, MAPE: 10.52%, Accuracy:89.48
(choy, 2021)	2021	Bursa Malaysia	Jan 2019 – Dec 2019	OHLCV	ARIMA, LSTM, SARIMA, Holt Winter, FB Prophet	RMSE, MAE	ARIMA RMSE: 0.0270 ARIMA MAE: 0.0189
(winata, 2021)	2021	Apple Stock	2010-2020	Close Price	SARIMAX	MAPE, RMSE	RMSE: 112.21, MAPE: 36.05%
(Garlapati, 2021)	2021	Apple Stock	2012-2021	Close Price	ARIMA, FBProphet	MAPE, RMSE	RMSE: 110.89, MAPE: 32.13%
(gao, 2021)	2021	Individual Stocks	2006-2017	OHLCV	Seq2Seq	MAPE, RMSE	-
(Qin, 2021)	2021	JD, PDD	2018-2021	OHLCV	ARIMA, BP	Relative Error	JD-ARIMA: 0.01047, JD-BPNN: 0.01219
(Wijesinghe, 2020)	2020	Colombo Stock Exchange	2017-2018	OHLCV	ARIMA, BP	MAE, MAPE	MAPE: 0.1783333, MAE: 29.6975
(majumder, 2019)	2019	Dhaka Stoc k Exchange	2010 – 2012	OHLCV	ARIMA	MAE, MAPE	MAPE: 17.27
(du, 2018)	2018	Shanghai Composite Index	2015 – 2016	OHLCV	ARIMA, BPNN	Accuracy	Accuracy: 78.79%
(xiong, 2017)	2017	Individual Stocks	2015-2016	OHLCV	ARIMA+BPNN	RMSE, MRE	RMSE: 18.02% MRE: 1.03%
(wichaidit, 2015)	2015	Stock Exchange of Thailand	2012-2015	OHLCV	CARIMA	MAE	MAE: 0.51
(li, 2015)	2015	SSE Composite Index	1990-2015	OHLCV	ARIMA	R-Square	R-Square: 0.0636

4.3. Studies Analyzing the Performance of Sentiment Analysis of Stock News Across the Globe

In the paper (Balshetwar and RS, 2023), the researchers introduced a methodology for detecting false news through the utilization of social media and news material, employing multiple imputations. The proposed approach involves the identification of pertinent feature terms through analyzing parts of speech (POS) in textual material. Addition- ally, sentiment analysis is utilized to estimate users' opinion control variables by employing lexicon-based scoring analysis. The research conducted revealed that the proposed approach exhibits a precision rate of 99.8% and an F1 Score of 95.6% in its ability to identify fabricated news.

In the paper (Ayyappa and Siva Kumar, 2022), the researchers present the utilization of Harris Hawks Induced Sparrow Search Optimization (HHISSO) for enhancing the efficacy of Long Short-Term Memory (LSTM) training through the manipulation of weights. Researchers employed the utilization of big data to store vast quantities of data, which was afterward analyzed to determine sentiment and make predictions on the stock market. The researchers employed cosine similarity as a computational technique following the data collection process. Term Frequency-Inverse Document Frequency (TFIDF) and Bag-of-Words (BOW) techniques are commonly employed in preparing news material. In the context of stock historical data, Exponential Moving Average (EMA), Average True Range (ATR), and True Range (TR) are frequently utilized. After implementing an optimized Long Short-Term Memory (LSTM) model for sentiment analysis using the Hybrid Hierarchical Input-Sequence Sentiment Score Optimization (HHISSO) technique, the predicted outcomes must be examined. Compared to alternative models, the proposed HHISSO + LSTM model demonstrates superior performance with a minimal error value of 0.018, suggesting a high degree of accuracy in predicting stock prices.

In the paper (Wu et al., 2021), the researchers introduced a sentiment model specifically designed for the stock market. The HAHTKN model is anticipated to yield financial gains for investors because to its ability to emulate human cognitive processes in the analysis of stock market data. The researchers collected a total of 3588 stock market communications from January to April 2019, sourced from the Anue website (https://www.cnyes.com/). In each epoch, the title is embedded, and a bidirectional GRU and attention approach are employed to obtain "vT" for each batch. Among the RNN cells examined, it was shown that 75% of the models had superior performance in the GRU cell, while the remaining 25% demonstrated better results in the LSTM cell.

In the paper (Li et al., 2021), the researchers proposed a sentiment analysis technique for Chinese stock reviews, utilizing BERT as the underlying model. The proposed model utilizes a BERT pre-training language model to encode stock reviews at the sentence level. The resulting feature vector is then passed to the classifier layer for classification. The findings indicate that the proposed technique performs better than TextCNN, TextRNN, Att-BLSTM, and TextCRNN in terms of precision, recall, and F1 scores. The BERT model, initially developed by Google, was utilized by researchers to undergo training using the Chinese Wikipedia data-set.

In the paper (Khan et al., 2020), the authors introduced a supervised approach that utilizes lexical, semantic, and a hybrid combination of both feature sets. The authors conducted the investigations on two separate occasions, employing both the complete message text and the specific spans as stated in Section. A com- parison was conducted on five classifiers: a decision tree classifier (Random Forest), a linear regression model (Linear Regression With SGD), a Lasso regression model, a Ridge regression model, and an SVR model. The initial four classifiers employed the MLlib implementation of Apache Spark, but the final classifier utilized Weka. Combining cosine similarity demonstrates optimally when BN synsets augment n-grams, increasing similarity scores from 0.663 to 0.677. In the data-set of micro-blog spans, the accuracy of f7 exhibits the maximum value when BN synsets and semantic frames are utilized to enhance n-grams, increasing from 0.712 to 0.726.

In the paper (Dridi et al., 2019), the authors employed SMS and email data to investigate the efficacy of learning classifier systems (LCS) rule-based machine learning techniques in the context of sentiment analysis about social media, movie reviews, and spam detection. The proposed classifier, XCSR#, demonstrates a substantial improvement compared to the baseline classifier.

In the paper (Nopp and Hanbury, 2015), the authors investigated the potential of sentiment analysis for banking supervisors in spotting risk. The online database Bankscope included information regarding the domicile, complete name, size as deter- mined by total assets, and Tier 1 capital ratio of banks after each year from 2001 to 2013. Risk sentiment analysis employs two independent methodologies. The experimental procedure consists of three sequential stages: (1) the acquisition and categorization of data, (2) the pre-processing of language and the identification of relevant features, and (3) the application of Naive Bayes (NB) and Support Vector Machine (SVM) classification techniques. The Sup- port Vector Machine (SVM) successfully classified around 92% of instances belonging to the UP class.

In the paper, (Mohapatra et al., 2012), the authors employ news and social media data as predictive indicators for stock market changes. The analysis consisted of a sample of 30 firms. Random Forest (RF) and Gradient Boosting Machine (GBM) have accurately predicted positive and negative future trend classes. Based on empirical investigations, it has been shown that social media and financial news exhibit the highest levels of forecast accuracy, specifically 80.53% and 75.16%, respectively.

In the paper (Schumaker et al., 2012), the researchers predicted financial news using Arizona Financial Text (AZFinText) and sentiment analysis. The study found that subjective news articles effectively predicted price direction with a success rate of 59.0%, compared to 50.0% for chance alone. Additionally, using subjective news items in a basic trading engine yielded a 3.30% return.

The summary of research studies based on stock sentiment analysis is shown in Table 4.

4.4. Studies Analyzing the Performance of AI Approaches on Various Stock Market Indices Across the Globe

In the paper (Srivastava et al., 2023), the researchers help customers analyze and speculate on stock price trends to buy stock for the most significant profit. The model employs LSTM and Twitter Tweets as Social Sentiment Indicators. Error calculation uses RMS (root mean square) error detection. The error percentage in the test data-set for each company is less than 1%.

Table 4 Summary of existing studies on stock price prediction using sentiment analysis.

Reference	Published Year	Task	Algorithm	Data Scope	Language	Performance
(Balshetwar, 2023)	2023	Sentiment classification	POS	News	English	Accuracy: 99.8% , F1 Score: 95.6%
(Deveikyte, 2022)	2022	Sentiment classification	VADER + LDA	News	English	Accuracy: 65%
(figa, 2022)	2022	Sentiment analysis	PCA	Tweets, Financial news	English	Accuracy: 72.00%
(ayyappa, 2022)	2022	Sentiment analysis	LSTM, HHISO, TFIDF	Financial news	English	Least Error Rate: 0.18
(Saxena, 2021)	2021	Sentiment analysis	LR, LDA, SVM, RF	Economic news	English	Accuracy: 90%
(wu, 2021)	2021	Sentiment analysis	GRU, LSTM	Anue website	English	Accuracy: 75%
(Das, 2022)	2021	Sentiment analysis	LSTM, SVC	Stock financial news, Tweets, Facebook comments	English	Accuracy: 98.11%
(Li, 2021)	2020	Sentiment analysis	RNN, LSTM, BERT	Economic news	English	Accuracy: 95%
(Kha, 2020)	2020	Sentiment analysis	RF, GBM	Twitter, Financial Times	English	Twitter Accuracy: 80.53% Financial News Accuracy: 75.16%
(dridi, 2019)	2019	Sentiment analysis	WEKA	Stockwits, Twitter	English	Cosine Similarity: 72.6%
(arif, 2018)	2018	Sentiment classification	XCRS#	Tweets, IMDB reviews	English	Accuracy: 75%
(li, 2017)	2017	Sentiment analysis	Backpropagation ANN	Financial news	English	-
(Khan, 2017)	2017	Sentiment analysis	TFIDF, MATLAB	Micro blogs	English	MSE: 0.5
(Dridi, 2016)	2016	Sentiment analysis	SVM, NB, Entropy	News	English	Accuracy: 99.65%
(nopp, 2015)	2015	Sentiment classification	SVM, NB	Online database Bankscope	English	Accuracy: 92%
(Mukherjee, 2014)	2014	Sentiment classification	-	Reuters	English	Accuracy: 86.04%
(Moraes, 2013)	2013	Sentiment classification	Fuzzy concept	Movies & E-books	English	Accuracy: 86.5%
(Schumaker, 2012)	2012	Sentiment analysis	VADER	Economic news	English	Accuracy: 50.0%
(Bai, 2011)	2011	Sentiment classification	SVM, NB	Camera support data	English	Accuracy: 92.70%

In the paper (Kamara et al., 2022), the researchers constructed a hybrid deep-learning network to estimate stock prices using long-term, short-term, and highly long-term data. For predicting stock values using CNNs, LSTMs, and attention to Johnson & Johnson and Bank of America investing data, researchers include Open, High, Low, Closing, Adjusted Close, and Volume of the data set. Mean Square Error, Mean Absolute Error, Mean Absolute Percentage Error, and Time Complexity were chosen for model evaluation. Adam optimizer was used to train the proposed model, where researchers claimed to train the model with 0.003, 1e-6 decay, and 0.9 momentum. In the model, the attention mechanism has an 8-neuron layer. The first layer of the ClusterBased LSTM (CB-LSTM) arm includes ten neurons, the second 30, and the third 20, including dropout layers, after the CBLSTM module's last two layers to prevent over-fitting and save computational power. The prediction unit has four levels. The output layer forecasts last. The first three layers have 20, 30, and 40 neurons. The strategy is tested with four optimizers and a fixed number of CNN layers. Optimization algorithms include Adam, Stochastic Gradient Descent (SGD), AdaGrad, and RMSprop. Researchers tested ABCNN with two, three, or four CNN layers per optimizer. The Adam optimizer with three CNN layers in the AB-CNN module has the lowest training and testing errors (0.047 and 0.039). With an 8.31% Sharpe Ratio (SR) and 77.34% Daily Winning Rate, PBMT has the most consistent returns. The researchers acknowledged certain limitations associated with the suggested technique, notably the requirement for a graphics processing unit (GPU) for practical training due to its inherent complexity. One other disadvantage is the need for extensive databases.

In the paper (Singh and Malhotra, 2023), the researchers employ a methodology that combines dual classifiers and sentiment analysis to accurately identify stock indexes. The integration of Decision Tree (DT) and Convolution Bi-Directional Gated Recurrent Unit (GRU) results in the formation of dual classifier networks. The network model under consideration is assessed based on the performance of Reliance Industries' shares. The accuracy of the sentiment analysis case study on Reliance Industries stock news was enhanced by 84.12% by the utilization of the adjusted sentiment index. Investor sentiment has a significant role in improving the accuracy of stock index trend prediction, as evidenced by a root mean square error (RMSE) of 3.16 and a decrease in the R-squared (R2) value of 0.97. The enhancement of predictive accuracy is substantially achieved through the adjustment of the sentiment index. The incorporation of investor sentiment inside the VADER (Value-Aware Dictionary and Sentiment Reasoner) and CNN + BDGRU models enhances the accuracy of stock price predictions for Reliance Industries.

In the paper (Ayala et al., 2021), the researchers propose employing technical indicators such as the Triple Exponential Moving Average (TEMA) and Moving Average Convergence/Divergence (MACD) in conjunction with a hybrid machine-learning methodology encompassing the Linear Model (LM), Artificial Neural Network (ANN), Random Forests (RF), and Support Vector Regression (SVR). Academic researchers gather daily trading data from the Ibex35 (IBEX), DAX, and DJI indices. Additional error measures commonly used in regression analysis are Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), and symmetric Mean Absolute Percentage Error (sMAPE). The TEMA crossover and MACD indicators are employed in the analysis. The execution of test "buy" and "sell" signals occurs on the subsequent trading day. Both trading strategies demonstrate that the Dow Jones Industrial Average (DJIA) has superior performance compared to the Deutsche Aktienindex (DAX) and the ´Indice Burs´atil Espa˜nol (IBEX). When applied to the IBEX and DAX indices, the Moving Average Convergence Divergence (MACD) indicator demonstrates superior performance to the Triple Exponential Moving Average (TEMA) indicator. However, its performance is comparatively worse when employed on the DJI index. TEMA exhibits a consistent pattern of underperformance on the IBEX index, frequently displaying a tendency to fall below the value of 1. The MACD indicator demonstrates efficacy in its functionality. However, it is essential to note that the profitability of the test data may be compromised due to tax implications. Additionally, the anticipated return does not sufficiently justify the associated risk. The predicted return of DAX is worthwhile solely when using the Moving Average Convergence Divergence (MACD) indicator. The DJI results for the TEMA, with a PF value of 4, exhibit a competitive nature.

In the paper (JuHyok et al., 2020), LSTM-based stock price reversal point prediction utilizing upward/downward feature sets was proposed. Twenty-seven feature candidates were developed from candlestick and technical indicators, and each stock's URP/DRP forecast feature sets were collected. LSTM-based URP/DRP predictors were then developed to improve accuracy. This model predicted 10 Chinese and 10 American stock reversals. The mean prediction accuracy (F1) for Chinese and American stock markets was 68.6% and 55.2%, respectively. The average prediction accuracy in China is 13.4% higher than in the US market. Compared to SVM, MLP, and CNN models, the proposed model increases F1 by 5.9%, 11.7%, and 5.3%.

In the paper (Fazeli and Houghten, 2019), the researchers examined prediction accuracy and evaluated the results. Historical stock data is sourced from an S&P 500 company. Our test data size was 120 days; thus, the researchers chose a stock with a trend change. The experiments used "Yahoo! Finance" data from 3/13/2014 to 3/12/2019 for Apple Stock (AAPL). LSTM networks form four layers. Features are resized, and sliding windows are built to produce input data. Data is transformed for LSTM network input. Our experiment without technical markers has a 0.04057 mean squared error.

In the paper (Vargas et al., 2018), the researchers build deep learning models to predict stock price daily using financial news and technical markers. Two models use distinct executions of indicators to predict daily directional movements of Chevron Corporation's (CVX) stock price. This study uses 106.494 Reuter's website news articles from October 20, 2006 to November 21, 2013. Experiments use Yahoo Finance's CVX stock price series. According to the results, accuracy exceeding 50% does not ensure lucrative forecasts. SI-RCNN-2 has a 51.08% accuracy and a loss of $935.03.

In the paper (Dash and Dash, 2016), the researchers have built a novel decision-making system and corresponding rules for generating trading decisions and attribute their achievement to utilizing a "Computationally Efficient Functional Link Artificial Neural Network" (CEFLANN). The decision-tree system employs a CEFLANN network, which evaluates the non-linear correlation among a limited number of widely used technical indicators to generate a constant flow of trading signals from 0 to 1. It is compared to the performance of several well-known machine learning methods, such as the Support Vector Machine (SVM), the Naive Bayesian model, the K closest neighbor model (KNN), and the Decision Tree (DT) model, to figure out how well the suggested method works. This analysis utilizes the price values of the BSE SENSEX and the S&P 500 stock indexes during the years 2010 and 2014. In this study, the training phase involves utilizing 1000 patterns from the BSE data set, while the testing phase is conducted on a subset of 208 patterns. Similarly, for the S&P data set, the training phase employs 1000 patterns, and the testing phase is performed on a subset of 221 patterns. Using a 5- fold cross-validation methodology and the initial 1000 samples as the training data set, it is possible to evaluate the network's generalization capacity. The training data set is partitioned into five subsets, with four subsets allocated for training purposes and the remaining subset used for validation. Conclusions on the data set are derived by utilizing the average outcomes obtained from 20 separate tests. After completing the training and validation processes, the network is evaluated using a separate data set not included in the original data set. Researchers claimed that the suggested model does a better job of making money than the support vector machine (SVM), the naive Bayesian model (NBM), the K closest neighbor model (KNN), and the decision tree (DT) models based on our analysis of the experimental data.

In the paper (Tan et al., 2008), the authors suggest combining deep learning and high- frequency technical indicators for investment. The elastic net model (SDNN) improves this strategy to handle the problems. This study examines the statistical and trading significance of indicators in-sample and out-of-sample. They suggest using an elastic net model to lower predictive component dimensions, improving high-dimensional prediction performance. Deep learning and Sharpe-Optimized methods can help investors balance risk and profit. In Chinese stock market trials, the Sharpe-Optimized SDNN outperformed the linear strategy by over 75% annually. SDNN outperforms LR, RF, XGBoost, and LSTM, as expected.

The summary of research studies based on artificial intelligence are shown in Table 5

Table 5 Summary of studies where AI approaches used in stock price prediction.

Reference	Published Year	Market	Period	Attributes	Predictor	Models	Performance Metrics	Trading Strategy
(Srivastava, 2023)	2023	Indian	2009-2020	OHLCV	LSTM	MLP	RMSE	Yes
(huang, 2023)	2023	Taiwan	2019-2021	OHLCV + 2 Tis	LSTM, GA, SVM, LR, DT	MLP	Accuracy, RMSE	Yes
(Pattewar, 2022)	2022	Tunisian	2016-2017	OHLCV	ADF, PP	RNN	Recall	Yes
Ayyappa, 2022	2022	Indian	2019-2020	OHLCV	LSTM, RNN, SVM, RF, NB	MLP	RMSE	Yes
(jing, 2021)	2021	Chinese	2017-2019	OHLCV + 16 Tis	CNN, LSTM, RNN, SVM	MLP	MAPE	Yes
(Alonso, 2020)	2020	Crypto currencies	2018-2019	18 Tis	CNN + LSTM	CNN, MLP, RBFNN	Accuracy, Statistical vali-dations	Yes
(Juhyok, 2020)	2020	Chinese & Americans	1997-2018	16 Candle Patterns + 10 Tis	LSTM	SVM, MLP, CNN	Accuracy, Precision, Recall, F1 score	Yes
(Kelotra, 2020)	2020	Indian	2017-2018	OHLCV + 12 Tis	Convolutional LSTM	DL, ARMA, NARX	MSE, RMSE	-
(li, 2021)	2020	Chinese	2003-2008	OHLCV + TIs + News	LSTM	SVM	Accuracy, Precision, Recall, F1- Score	-
(Qiu, 2020)	2020	Chinese	2016-2019	9 Tis	LSTM + GRU	12 Models	Accuracy, Precision, Recall, F1- Score	-
(song, 2020)	2020	Korean	2000-2019	250 Binary Events	DNN	DNN with Tis	Accuracy	Yes
(Demir, 2019)	2019	Belgian	2014-2018	9 Tis	NN, CNN, ResNet	NN,CNN, ResNet	RMSE, MAE, PCC	-
(Wen, 2020)	2019	Chinese	2016-2018	OHLCV + Turnover	PCA + LSTM	CNN, MLP, MA	RMSE, MAPE	-
(Agrawal, 2019)	2019	Indian	2016-2018	OHLCV + 2Tis	Optimal LSTM	MLP, ELSTM, LR, SVM	Accuracy, MSE	Yes
(al,2019)	2019	Bahrain	2010-2018	Close +2 Tis	Auto LSTM: LSTM autoen-code	LSTM, MLP	MAE, RMSE	-
(Lai, 2019)	2019	Taiwanese	2009-2018	OHLCV + 4 Tis	ARIMA + LSTM	-	MAE, RMSE	-
(Borovkova, 2019)	2019	American	2014	OHLCV + 12 Tis	LSTM	Ridge regression, Lasso regression	AUC, ROC	Yes
(Chen, 2018)	2018	Chinese	2004-2018	OHLCV + 14 Tis	LSTM with Attention and	LSTM	MAE, MSE	-

					Market Vector			
(Labiad, 2018)	2018	Moroccan	2016-2017	21 TIs	LSTM	MLP	Accuracy, Precision, Recall, F1 Score	-
(Oncharoen, 2018)	2018	American	2006-2017	OHLCV + 7 TIs	CNN + LSTM	CNN + LSTM	Accuracy	Yes
(Wang, 2018)	2018	American	2002-2017	15 TIs	CNN + TA	RSI, MA, LSTM, MLP	Accuracy, Precision, Recall, F1- Score	Yes
(Vargas, 2017)	2018	American	2006-2013	7 TIs + News	CNN + LSTM	ANN	Accuracy	-
(Tan, 2020)	2018	American	2006-2013	6 TIs + News	CNN + LSTM	SI-RCNN, I- RNN	Accuracy	Yes
(Lee, 2017)	2018	American	2016-2018	OHLCV + 9 TIs + News	LSTM	LSTM	Accuracy, Statistical Vali-dations	-
(Nelson, 2017)	2018	Chinese	2012-2015	14 TIs + News	DBN	ANN, SVM, RF, DBN, RNN, LSTM	Accuracy, Precision, Recall, F1-Score, AUC	
(Das16)	2017	American	2016	Close price + 8 TIs	LSTM	MLP	RMSE	-

4.5. Classification of stock market decision-making processes

The researchers mentioned in the referenced article (Yang et al, 2022) focused their attention on financial and technical sentiment indicators. In this study, the researchers collected financial literature and stock price histories from online sources, afterwards quantifying sentiment and technical indicators. In contrast to the existing indicator method, this alternative approach offers a more comprehensive depiction of stock prices and trade sentiment. To infer customer sentiment from gigabytes of financial data, the authors employed the FinBERT model. The efficacy of this transfer learning model in financial sentiment analysis is enhanced by its pre-training on a substantial corpus of financial literature. The authors utilized Python's E-finance tools to extract stock data from three distinct companies for the period of 2012 to 2020. The historical transaction data set for stocks encompasses several parameters such as op prices, closing prices, highs, lows, and trading volume. The authors employed the Technical Trading Rules (TTR) software to ascertain the numerical values associated with 37 widely recognized technical indicators. The study describes the LASSO-LSTM model, which effectively addresses the issue of multi-collinearity and noise in predictor variables through its memory structure and variable selection capabilities. The task of forecasting changes in stock prices, specifically whether to buy or sell, can be characterized as a binary classification problem. The accuracy metric may have difficulties in accurately evaluating the predictive performance of a model when the distribution of the sample is not uniform. To evaluate the precision of the model's predictions, the researchers incorporated many novel measures, such as the Kappa coefficient, ROC curve, sensitivity, and specificity confusion matrix. The evaluation of a model's predictive performance can be assessed using metrics such as accuracy, sensitivity, and specificity, which can be obtained from the confusion matrix. In order to mitigate the influence of scale disparities among the predictor variables, a standardization process is employed within the LASSO-LSTM model. Subsequently, the coordinate algorithm is employed for the purpose of parameter allocation. In this study, the LASSO methodology is employed to choose the most significant variables from a pool of 41 indicators. Additionally, a 10-fold cross-validation technique and a grid search method are utilized to identify the optimal λ values. The process involves the identification of an optimal value for the regularization parameter λ and the application of the LASSO method to reduce the number of variables, utilizing the information obtained from the training data set. Once the appropriate variables have been meticulously chosen, the training set data is utilized to instruct a LASSO-LSTM model in forecasting stock movements, namely the decision to either buy or sell. Finally, compared the prediction performance of LASSO-LSTM, RF, LSTM, and SVM is based on four metrics.

5. Research Design

The systematic review criteria and literature selection procedures are covered here

5.1. Research Framework

Random sampling yielded 146 related essays. These comprise 2011–2023 journal publications, conference proceedings papers, PhD dissertations, and unpublished academic working papers and reports. The selected studies were classified into three major prediction approaches: technical analysis, statistical approach, sentiment analysis of stock news, and artificial intelligence approach. Second, the conceptual framework was used to investigate each group's unique qualities.

5.2. Distribution of literature

Table 6 shows the distribution of surveyed works based on categorization into the technical analysis, statistical approach, sentiment analysis of stock news, and Artificial intelligence approach. Thirty-three (33) were based on technical analysis, twelve (12) were based on a statistical approach, twenty-eight (28) were based on stock sentiment analysis, and thirty-two (32) were based on the artificial intelligence approach.

Table 6 Distribution of the literature based on categorization.

Category	Number of research studies	Percentage
Technical analysis	43	29.45%
Statistical approach	23	15.75%
Stock sentiment analysis	38	26.02%
Artificial Intelligence	42	28.76%

5.3. Technical Analysis

The study found that 29.45% of examined publications relied on technical analysis, as indicated in Table 8, due to the structure of past stock prices. The lowest data collection term was four years by Göçken et al. (2016), and the maximum was 12 years by Salles et al. (2018) in 2018. The simple moving average was the most used technical indicator. Tables 2 and 3 show that several Simple Moving Average (SMA) research used MVR, OHLCV, EMA, MACD, OBV, and RSI. A total of 27 research predicted share prices daily and 3-day periods. Interestingly, just 28% of the 146 publications studied focused on the Asian market. Tables 2 and 3 reveal that most studies predicted a day, followed by intra-day in the ranges of 1, 3, 5, 7, 8, 10 days, and 1-week to 2 months.

5.4. Statistical Analysis

The study reveals that 15.75% of surveyed works were based on statistical approaches such as ARIMA, SARIMA, and Facebook prophet, as shown in Table 4. Most studies that considered OHLCV parameters for model training used ARIMA model followed by Back-propagation and neural networks. A research study Kamara et al. (2022) published in the year 2022 where the authors claimed maximum accuracy 89.48% in the prediction of stock price whereas research study Choy et al. (2021) published in the year 2021 where researchers claimed least RMSE of 0.0270 and MAE of 0.0189 using ARIMA technique.

5.5. Stock News Sentiment Analysis

The study reveals that 26.02% of surveyed works were based on sentiment analysis of stock news, as shown in Table 8. Most studies are based on sentiment classification, followed by sentiment analysis. Sentiment classification revolves around identifying text polarities such as positive, neutral, and negative. Other studies also include five classes of polarities such as very positive, positive, neutral, negative, and very negative. Financial News and Twitter became the most trusted data sources mentioned in most studies. Research study Balshetwar (2023) published in the year 2023 where researchers claimed 99.8% accuracy and 95.6% F1-score in the sentiment classification task.

5.6. Artificial Intelligence

The study reveals that 28.76% of surveyed works were based on artificial intelligence in the prediction of share price, as shown in Table 8. The maximum time frame chosen for data collection was twenty-one (21) years in a research study (JuHyok et al., 2020) published in the year 2020, and the minimum time frame was one year in a research study (Yang et al 2022) published in the year 2022. The deep learning model Long-Short Term Memory (LSTM) is one of the most used models (17 studies) used by researchers in their studies to predict stock price movement and its price.

5.7. Discussion and Analysis

This section provides a concise overview of the latest and advanced algorithms identified in the reviewed articles. Consequently, the primary objective entailed comprehensive documentation of all utilized materials and a thorough elucidation of any omissions, so facilitating a contextual framework for subsequent scholarly inquiry. The task was successfully completed by employing the use of the. The research findings indicated that a significant majority, specifically 87 percent, of the implemented AI algorithms consisted of technical signals. This shows that the bulk of the research that was analysed anticipated changes in the price of the stock. There is a scarcity of magazines available that provide prognostications regarding future stock prices. Hence, a prospective area of investigation in the future may revolve around the challenges associated with predicting both the financial implications and the dynamics involved.

5.8. Investigation 1: Machine learning and deep learning approaches have been widely employed to forecast the direction, movement, and price of shares, as they are the most prevalent methods in this domain.

Figure 10 illustrates the distribution of publications based on different prediction models. The research conducted in this study mostly centered on the prediction of share price utilizing Machine and Deep learning techniques. This paper comprehensively analyzes extensively researched models for predicting share prices and a detailed chronological examination of prediction studies conducted over the years.

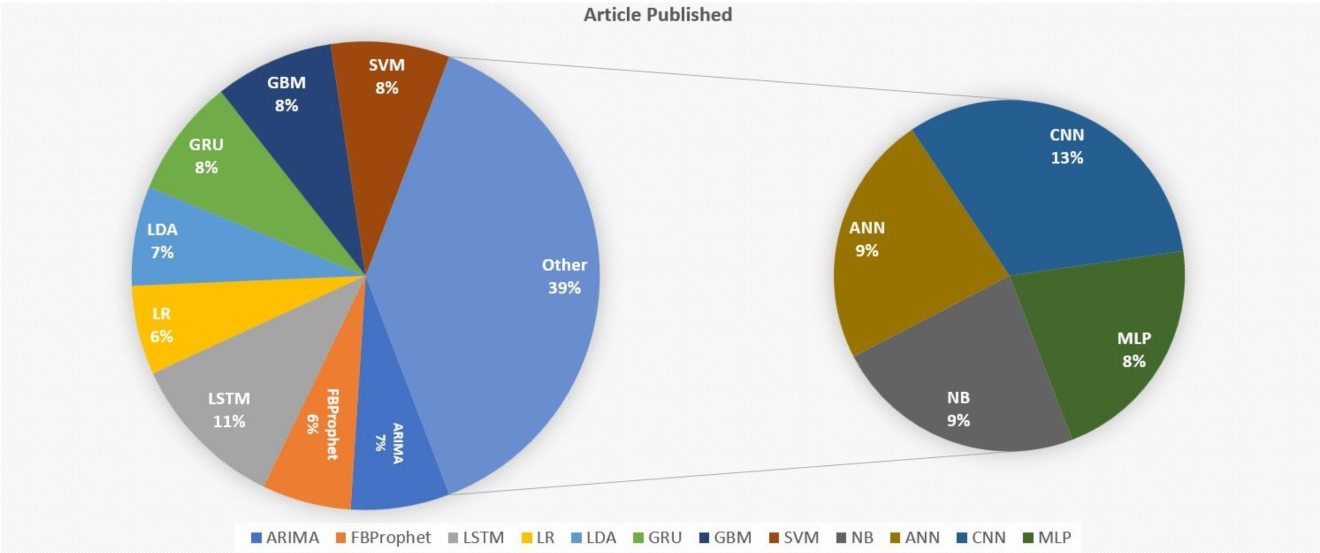

Figure 10 Research studies published on basis of prediction models.

Again, ARIMA, LSTM, and CNN are the most popular AI algorithms for stock market predictions, followed by ANN and SVM. Over 50% of the assessed works utilized hybrid algorithms to address weaknesses in separate algorithms, resulting in higher accuracy compared to other models of the same type. Thus, studying hybrid algorithms in stock-market prediction may yield fresh insights that intrigue future academics.

In Figure 11, it has been observed that most of the researchers used Simple Moving Average (SMA) as a technical indicator followed by Relative Strength Index (RSI) and OHLCV.

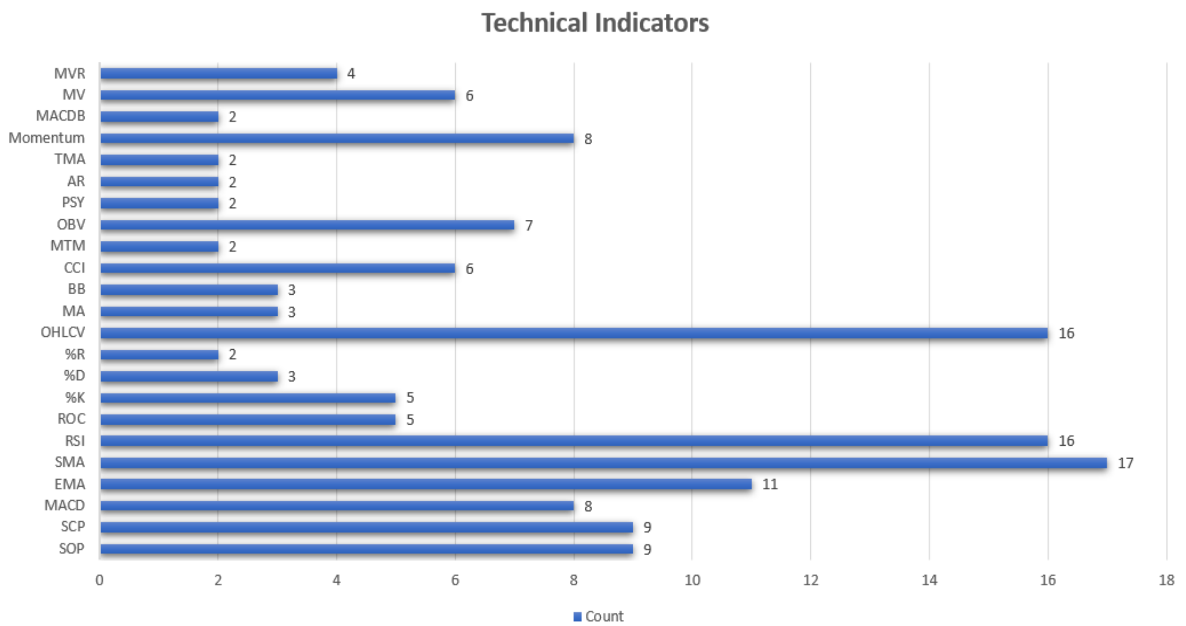

Figure 11 Technical indicators used by researchers in stock price forecast.

5.9. Investigation 2: *The widely used share market data sites and year when most share price prediction studies have been published*

The PRISMA criteria find machine learning and deep learning research articles forecasting share prices.

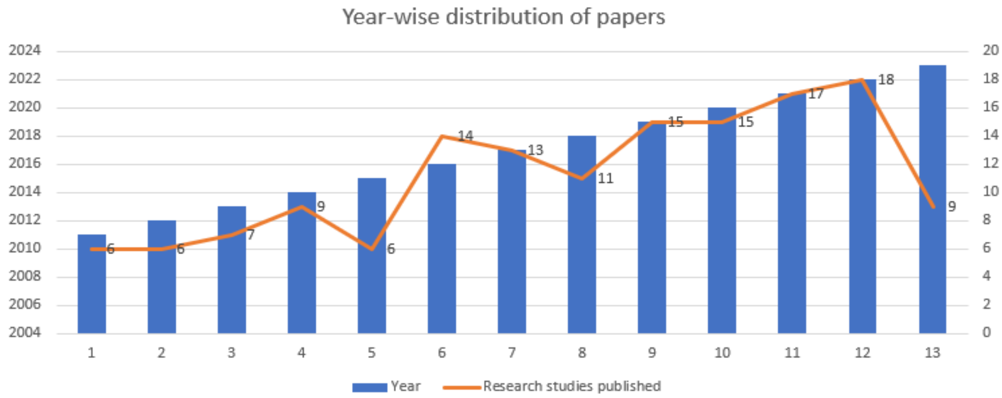

Figure 12 Research studies published year wise since year 2011.

The chosen research papers' year distribution is displayed in Fig. 12. The time frame was 2011–2023. Research articles on share price prediction peaked in 2022 with 18 articles, while the lowest number was observed in 2011-2015 with six articles.

5.10. Investigation 3: The metrics are used to evaluate a machine and deep learning model

The evaluation and assessment of prediction models are paramount in determining their efficacy. Table 7 presents a concise overview of the evaluation indicators associated with the different prediction models. Accuracy (Acc), Sensitivity/Recall (Re), Specificity (S_p), and Precision (Pr) are often employed measures for quantifying categorization performance in academic research publications (Ingle, 2021). Additional performance evaluation metrics, including F-measure (F), R2, RMSE, MAE, and MAPE, have been employed in forecasting jobs to assess predictive accuracy (Srivastava et al., 2023). The research conducted by (Carta et al., 2021) indicates that accuracy is the primary criterion for evaluating overall performance. Furthermore, specificity is a statistical measure that estimates the proportion of actual negative observations relative to an algorithm's anticipated number of negatives. This ratio is called the "ratio" of true negative observations to expected negatives."

In Figure 13, it has been observed that Accuracy parameters used at the most by researchers to evaluate their model followed by Root Mean Square Error (RMSE) and Magnitude of Relative Error (MRE). This concept is often called "actual negative rate". Sensitivity is an algorithm's ability to anticipate true positive outcomes from positive data. The F1 score is a mathematical combination of accuracy and recall measures that specifically uses the weighted harmonic mean. Comparing optimistic and negative forecasts determines model accuracy. The performance indication above is more reliable than accuracy and precision. Analyzing heavily skewed data with the latter measurements can lead to incorrect interpretations. To evaluate a classification model using test data, a high F1 score is best. This method helps classification models balance accuracy and recall. The state-of-art evaluation parameters used in evaluate the results of stock prediction framework mentioned in the following table 7.

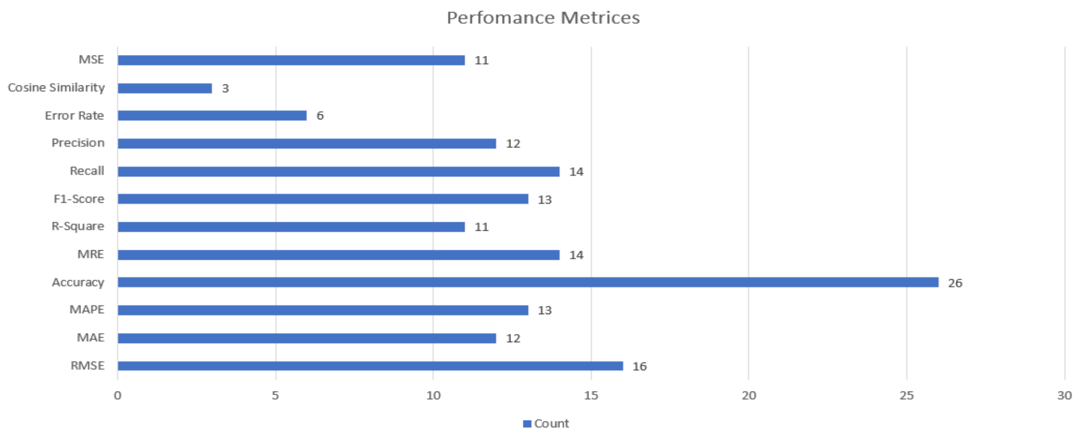

Figure 13 Use of performance metrics used by researchers since 2011 for stock price forecast.

Table 7 Summary of model evaluation metrices.

Parameter	Description	Formula	Articles
Accuracy	It is the number of examples where the model achieved perfect recognition.	$\frac{TR+TF}{TR+FR+FF+TF}*100$	(Wang Z. a.-B., 2023)
Precision	A prediction's precision is the percentage of stock gains or losses that actually occur	$\frac{TR}{TR+FR}*100$	(Silva, 2020)
Recall	This value is a percentage of the model's expected stock price rise or decline.	$\frac{TR}{TR+FF}*100$	(Srivastava, 2023)
F1 Score	When a predictor's precision and recall are equal, F1-Score links the correct stock (rise/fall) to the predicted stock	$\frac{2*P^R*R^R}{P^R+R^R}$	(Ayala, 2021)
R^2	From 0 to 1, the performance index shows the correlation between anticipated and actual values.	$\frac{\sum_{i=1}^{n}(t_i-t)(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(t_i-\bar{t})^2.\sum_{i=1}^{n}(y_i-\bar{y})^2}}$	(Dridi, 2019)
MAE	This measure calculates the average magnitude of all data set predictions without considering error direction.	$\frac{1}{n}\sum_{i=1}^{n}(t_i-y_i)$	(Haq, 2021)
MAPE	The lower MAPE variable is better for measuring average absolute percentage error than the higher one.	$\frac{1}{n}\sum_{i=1}^{n}\left\Vert\frac{t_i-y_i}{t_i}\right\Vert$	(Ji, 2023)
Gross Loss	Nevertheless, a loss might be considered the opposite of a gain. Investors incur losses when the market value of their investments decreases below the price at which they were initially acquired.	$= Netsales - COGS$	(Jing, 2021)
Gross Profit	The gross profit of a firm refers to the total revenue generated by the company before subtracting any costs associated with the manufacturing and distribution of its products or the provision of its services. Determining a company's gross profit, as presented in the income statement, involves deducting the cost of goods sold (COGS) from the generated sales revenue. The financial figures in question can be seen within the income statement of a corporation.	$= Netsales - COGS$	(Nopp, 2015)
Rate of Return	The rate of return (RoR) refers to the financial gain or loss experienced on an investment about its initial value throughout a specific duration. The rate of return is defined as the percentage difference between the initial and final values after a given period.	$\frac{Current\ value - Initial\ value}{Initial\ value}$	(Zhang L. a., 2019)
Sharpe Ratio	The Sharpe ratio assesses the profitability of an investment relative to its associated level of risk. The presence of excess returns within a specific time frame may suggest a rise in volatility and risk rather than demonstrating investment proficiency. In equation, $R_p$is *return of portfolio*, R_f is *risk-free rate* and σ_p is *standard deviation of portfolio's excess return*.	$\frac{R_p-R_f}{\sigma_p}$	(Picasso, 2019)
Average Coefficient of Variation	he statistical measure used to quantify the dispersion of a data series around its mean is known as the coefficient of variation (CV). The coefficient of variation can be employed to assess the extent of variance between two data sets, even if their means exhibit significant disparity. The ratio of the standard deviation to the mean is a fundamental metric. In equation, σ_p is standard deviation and μ is mean.	$=\frac{\sigma}{\mu}$	(Shahvaroughi Farahani, 2021)

6. Final considerations

In this article, the authors aimed to survey computational intelligence's important and up-to-date contributions to solve stock market forecasting problems. This paper breaks down Stock market forecasting strategies into four broad categories: prediction techniques using technical analysis, statistical techniques, sentiment analysis, and artificial intelligence. Using 146 research articles, said survey aims to categorize existing strategies according to the years in which studies were first published, the methodologies adopted in studies, the data sets that researcher used, performance indicators, and implementation tools. ANN, SVM, SVR, HMM, NN, fuzzy-based approaches, LSTM, XGBoost, K-means, and so on are some techniques used for stock market prediction. For this purpose, the research gaps and the challenges related to predicting the stock market are discussed. This survey shows that CNN and LSTM are two of the most regularly utilized techniques for predicting the stock market. Stock market regulation and monitoring can be accomplished using these strategies. The significant outcomes of our research are: (1) The utilization of technical indicators in predicting stock market trends is a widely adopted activity. The issue of effectively identifying a set of technical indicators remains unresolved. (2) Using suitable pre-processing and feature selection procedures can enhance the effectiveness of stock market forecasting models. (3) Furthermore, the issue of accurately predicting stock market trends lends itself well to the application of computational intelligence technologies. Hybrid models, which incorporate the predictive capabilities of distinct base models, are progressively gaining traction in stock market prediction. (4) Furthermore, it is worth noting that diverse authors have utilized numerous assortments of performance measurements, hence precluding the identification of a singular metric that can be deemed as conclusive. The insufficiency of existing methodologies for using historical stock data poses a significant obstacle to the stock price forecasting system. Several factors, including governmental choices and market attitudes, contribute to the challenges of effectively identifying these practices. Hence, acquiring information from many sources is imperative to developing well-founded assessments, and data pre-processing presents several challenges for individuals involved in data mining. Given the constraints above, anticipating future trends in the stock market poses significant difficulties, necessitating the utilization of state-of-the-art methodologies. To address these substantial constraints, investors must have access to cutting-edge techniques for predicting stock market trends.

Ethical considerations

This material is the authors' own original work, which has not been previously published elsewhere

Conflict of 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.

Funding

The authors did not receive support from any organization for the submitted work.

References

Abhishek, K., Khairwa, A., Pratap, T., & Prakash, S. (2012, July). A stock market prediction model using Artificial Neural Network. In *2012 Third International Conference on Computing, Communication and Networking Technologies (ICCCNT'12)* (pp. 1-5). IEEE.

Agrawal, M., Shukla, P. K., Nair, R., Nayyar, A., & Masud, M. (2022). Stock Prediction Based on Technical Indicators Using Deep Learning Model. *Computers, Materials & Continua*, *70*(1).

Agrawal, A., Mohanty, P., & Totala, N. K. (2019). Does EVA beat ROA and ROE in explaining the stock returns in Indian scenario? An evidence using mixed effects panel data regression model. *Management and Labour Studies*, *44*(2), 103-134.

Argiddi, R. V., & Apte, S. S. (2012). Future trend prediction of Indian IT stock market using association rule mining of transaction data. *International Journal of Computer Applications*, *39*(10), 30-34.

ARIF, M. F. (2018). The Factors Affecting Voluntary Auditor Switching in Companies Listed in Indonesia Stock Exchange. *Qualitative and Quantitative Research Review*, *3*(2).

Atsalakis, G. S., Dimitrakakis, E. M., & Zopounidis, C. D. (2011). Elliott Wave Theory and neuro-fuzzy systems, in stock market prediction: The WASP system. *Expert Systems with Applications*, *38*(8), 9196-9206.

Ayala, J., García-Torres, M., Noguera, J. L. V., Gómez-Vela, F., & Divina, F. (2021). Technical analysis strategy optimization using a machine learning approach in stock market indices. *Knowledge-Based Systems*, *225*, 107119.

Ayyappa, Y., & Siva Kumar, A. P. (2022). Optimized long short-term memory-based stock price prediction with sentiment score. *Social Network Analysis and Mining*, *13*(1), 13.

Balshetwar, S. V., & Rs, A. (2023). Fake news detection in social media based on sentiment analysis using classifier techniques. *Multimedia Tools and Applications*, 1-31.

Bao, D., & Yang, Z. (2008). Intelligent stock trading system by turning point confirming and probabilistic reasoning. *Expert Systems with Applications*, *34*(1), 620-627.

Bai, X. (2011). Predicting consumer sentiments from online text. *Decision Support Systems*, *50*(4), 732-742.

Bisoi, R., & Dash, P. K. (2014). A hybrid evolutionary dynamic neural network for stock market trend analysis and prediction using unscented Kalman filter. *Applied Soft Computing*, *19*, 41-56.

Borovkova, S., & Tsiamas, I. (2019). An ensemble of LSTM neural networks for high-frequency stock market classification. *Journal of Forecasting*, *38*(6), 600-619.

Bustos, O., Pomares, A., & Gonzalez, E. (2017, October). A comparison between SVM and multilayer perceptron in predicting an emerging financial market: Colombian stock market. In *2017 Congreso Internacional de Innovacion y Tendencias en Ingenieria (CONIITI)* (pp. 1-6). IEEE.

Carta, S., Ferreira, A., Podda, A. S., Recupero, D. R., & Sanna, A. (2021). Multi-DQN: An ensemble of Deep Q-learning agents for stock market forecasting. *Expert systems with applications*, *164*, 113820.

Chandar, S. K. (2019). Stock market prediction using subtractive clustering for a neuro fuzzy hybrid approach. *Cluster Computing*, *22*(Suppl 6), 13159-13166.

Chandar, S. K. (2022). Convolutional neural network for stock trading using technical indicators. *Automated Software Engineering*, *29*, 1-14.

Che, L. (2018). Investor types and stock return volatility. *Journal of Empirical Finance*, *47*, 139-161.

Chen, Y., & Hao, Y. (2017). A feature weighted support vector machine and K-nearest neighbor algorithm for stock market indices prediction. *Expert Systems with Applications*, *80*, 340-355.

Cheng, W. X., Suganthan, P. N., Qiu, X., & Katuwal, R. (2020). Classification of stock market trends with confidence-based selective predictions. In *Swarm, Evolutionary, and Memetic Computing and Fuzzy and Neural Computing: 7th International Conference, SEMCCO 2019, and 5th International Conference, FANCCO 2019, Maribor, Slovenia, July 10–12, 2019, Revised Selected Papers 7* (pp. 93-104). Springer International Publishing.

Choy, Y. T., Hoo, M. H., & Khor, K. C. (2021, September). Price prediction using time-series algorithms for stocks listed on bursa malaysia. In *2021 2nd International Conference on Artificial Intelligence and Data Sciences (AiDAS)* (pp. 1-5). IEEE.

Das, S. P., & Padhy, S. (2012). Support vector machines for prediction of futures prices in Indian stock market. *International Journal of Computer Applications*, *41*(3).

Das, N., Sadhukhan, B., Chatterjee, T., & Chakrabarti, S. (2022). Effect of public sentiment on stock market movement prediction during the COVID-19 outbreak. *Social network analysis and mining*, *12*(1), 92.

Dash, R., & Dash, P. K. (2016). A hybrid stock trading framework integrating technical analysis with machine learning techniques. *The Journal of Finance and Data Science*, *2*(1), 42-57.

Demir, C. (2019). Macroeconomic determinants of stock market fluctuations: The case of BIST-100. *Economies*, *7*(1), 8.

Deveikyte, J., Geman, H., Piccari, C., & Provetti, A. (2022). A sentiment analysis approach to the prediction of market volatility. *Frontiers in Artificial Intelligence*, *5*, 836809.

Di Persio, L., & Honchar, O. (2017). Recurrent neural networks approach to the financial forecast of Google assets. *International journal of Mathematics and Computers in simulation*, *11*, 7-13.

Dinesh, S., Rao, N., Anusha, S. P., & Samhitha, R. (2021, April). Prediction of trends in stock market using moving averages and machine learning. In *2021 6th International Conference for Convergence in Technology (I2CT)* (pp. 1-5). IEEE.

Dospinescu, N., & Dospinescu, O. (2019). A Profitability Regression Model in Financial Communication of Romanian Stock Exchange S COMPANIES. *Ecoforum Journal*, *8*(1).

Dridi, A., Atzeni, M., & Reforgiato Recupero, D. (2019). FineNews: fine-grained semantic sentiment analysis on financial microblogs and news. *International Journal of Machine Learning and Cybernetics*, *10*, 2199-2207.

Enke, D., & Mehdiyev, N. (2013). Stock Market Prediction Using a Combination of Stepwise Regression Analysis, Differential Evolution-based Fuzzy Clustering, and a Fuzzy Inference Neural Network. *Intelligent Automation & Soft Computing*, *19*(4).

Enke, D., Grauer, M., & Mehdiyev, N. (2011). Stock market prediction with multiple regression, fuzzy type-2 clustering and neural networks. *Procedia Computer Science*, *6*, 201-206.

Fama, E. F. (1998). Market efficiency, long-term returns, and behavioral finance. *Journal of financial economics*, *49*(3), 283-306.

Fang, Y., Fataliyev, K., Wang, L., Fu, X., & Wang, Y. (2014, July). Improving the genetic-algorithm-optimized wavelet neural network for stock market prediction. In *2014 International Joint Conference on Neural Networks (IJCNN)* (pp. 3038-3042). IEEE.

Fazeli, A., & Houghten, S. (2019, December). Deep learning for the prediction of stock market trends. In *2019 IEEE International Conference on Big Data (Big Data)* (pp. 5513-5521). IEEE.

Figà-Talamanca, G., & Patacca, M. (2022). An explorative analysis of sentiment impact on S&P 500 components returns, volatility and downside risk. *Annals of Operations Research*, 1-23.

Gaba, S., Gupta, M., & Singh, H. (2023, October). A comprehensive survey on VANET security attacks. In *AIP Conference Proceedings* (Vol. 2495, No. 1). AIP Publishing.

Gao, R., Zhao, Y., & Zhang, B. (2021). The spillover effects of economic policy uncertainty on the oil, gold, and stock markets: Evidence from China. *International Journal of Finance & Economics*, *26*(2), 2134-2141.

Gao, S. E., Lin, B. S., & Wang, C. M. (2018, December). Share price trend prediction using CRNN with LSTM structure. In *2018 International Symposium on Computer, Consumer and Control (IS3C)* (pp. 10-13). IEEE

García-Medina, A., Junior, L. S., Bañuelos, E. U., & Martínez-Argüello, A. M. (2018). Correlations and flow of information between the New York Times and stock markets. *Physica A: Statistical Mechanics and its Applications*, *502*, 403-415.

Garlapati, A., Krishna, D. R., Garlapati, K., Rahul, U., & Narayanan, G. (2021, April). Stock price prediction using Facebook Prophet and Arima models. In *2021 6th International Conference for Convergence in Technology (I2CT)* (pp. 1-7). IEEE.

Gite, S., Patil, S., Dharrao, D., Yadav, M., Basak, S., Rajendran, A., & Kotecha, K. (2023). Textual feature extraction using ant colony optimization for hate speech classification. *Big data and cognitive computing*, *7*(1), 45.

Göçken, M., Özçalıcı, M., Boru, A., & Dosdoğru, A. T. (2016). Integrating metaheuristics and artificial neural networks for improved stock price prediction. *Expert Systems with Applications*, *44*, 320-331.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.

Grigoryan, H. (2017, May). Stock market trend prediction using support vector machines and variable selection methods. In *2017 International Conference on Applied Mathematics, Modelling and Statistics Application (AMMSA 2017)* (pp. 210-213). Atlantis Press.

Gupta, D., Anand, A., & Singh, R. (2008). Empirical testing of strong form of market efficiency. *Asia Pacific Business Review*, *4*(3), 53-59.

Gururaj, V., Shriya, V. R., & Ashwini, K. (2019). Stock market prediction using linear regression and support vector machines. *Int J Appl Eng Res*, *14*(8), 1931-1934.

Haq, A. U., Zeb, A., Lei, Z., & Zhang, D. (2021). Forecasting daily stock trend using multi-filter feature selection and deep learning. *Expert Systems with Applications*, *168*, 114444.

Huang, Y., Ma, F., Bouri, E., & Huang, D. (2023). A comprehensive investigation on the predictive power of economic policy uncertainty from non-US countries for US stock market returns. *International Review of Financial Analysis*, *87*, 102656.

Idrees, S. M., Alam, M. A., & Agarwal, P. (2019). A prediction approach for stock market volatility based on time series data. *IEEE Access*, *7*, 17287-17298.

Ingle, V., & Deshmukh, S. (2021). Ensemble deep learning framework for stock market data prediction (EDLF-DP). *Global Transitions Proceedings*, *2*(1), 47-66.

Ji, Z., Wu, P., Ling, C., & Zhu, P. (2023). Exploring the impact of investor's sentiment tendency in varying input window length for stock price prediction. *Multimedia Tools and Applications*, 1-35.

Jing, N., Wu, Z., & Wang, H. (2021). A hybrid model integrating deep learning with investor sentiment analysis for stock price prediction. *Expert Systems with Applications*, *178*, 115019.

Joseph, K., Wintoki, M. B., & Zhang, Z. (2011). Forecasting abnormal stock returns and trading volume using investor sentiment: Evidence from online search. *International Journal of Forecasting*, *27*(4), 1116-1127.

JuHyok, U., Lu, P., Kim, C., Ryu, U., & Pak, K. (2020). A new LSTM based reversal point prediction method using upward/downward reversal point feature sets. *Chaos, Solitons & Fractals*, *132*, 109559.

Kaczmarek, T., & Perez, K. (2022). Building portfolios based on machine learning predictions. *Economic research-Ekonomska istraživanja*, *35*(1), 19-37.

Kamara, A. F., Chen, E., & Pan, Z. (2022). An ensemble of a boosted hybrid of deep learning models and technical analysis for forecasting stock prices. *Information Sciences*, *594*, 1-19.

Kelotra, A., & Pandey, P. (2020). Stock market prediction using optimized deep-convlstm model. *Big Data*, *8*(1), 5-24.

Khan, W., Malik, U., Ghazanfar, M. A., Azam, M. A., Alyoubi, K. H., & Alfakeeh, A. S. (2020). Predicting stock market trends using machine learning algorithms via public sentiment and political situation analysis. *Soft Computing*, *24*, 11019-11043.

Khan, W., Ghazanfar, M. A., Azam, M. A., Karami, A., Alyoubi, K. H., & Alfakeeh, A. S. (2020). Stock market prediction using machine learning classifiers and social media, news. *Journal of Ambient Intelligence and Humanized Computing*, 1-24.

Khan, T. A. (2011). Cointegration of international stock markets: An investigation of diversification opportunities. *Undergraduate Economic Review*, *8*(1), 7.

Kim, S., Ku, S., Chang, W., & Song, J. W. (2020). Predicting the direction of US stock prices using effective transfer entropy and machine learning techniques. *IEEE Access*, *8*, 111660-111682.

Kumar, D. A., & Murugan, S. (2013, February). Performance analysis of Indian stock market index using neural network time series model. In *2013 international conference on pattern recognition, informatics and mobile engineering* (pp. 72-78). IEEE.

Kumar, G., Jain, S., & Singh, U. P. (2021, March). Neuro-fuzzy and particle swarm optimization based hybrid approach for stock price forecasting. In *2021 International Conference on Emerging Smart Computing and Informatics (ESCI)* (pp. 753-758). IEEE.

Kumar Naik, P., Shaikh, I., & Duc Huynh, T. L. (2022). Institutional investment activities and stock market volatility amid COVID-19 in India. *Economic research-Ekonomska istraživanja*, *35*(1), 1542-1560.

Labiad, B., Berrado, A., & Benabbou, L. (2018, October). Short term prediction framework for moroccan stock market using artificial neural networks. In *Proceedings of the 12th International Conference on Intelligent Systems: Theories and Applications* (pp. 1-6).

Lee, U. (1992). Do stock prices follow random walk?:: Some international evidence. *International Review of Economics & Finance*, *1*(4), 315-327.

Li, M., Li, W., Wang, F., Jia, X., & Rui, G. (2021). Applying BERT to analyze investor sentiment in stock market. *Neural Computing and Applications*, *33*, 4663-4676.

Li, J., Bu, H., & Wu, J. (2017, June). Sentiment-aware stock market prediction: A deep learning method. In *2017 international conference on service systems and service management* (pp. 1-6). IEEE.

Li, Y., & Giles, D. E. (2015). Modelling volatility spillover effects between developed stock markets and Asian emerging stock markets. *International Journal of Finance & Economics*, *20*(2), 155-177.

Lin, Y. F., Ueng, Y. L., Chung, W. H., & Huang, T. M. (2019, May). Stock price range forecast via a recurrent neural network based on the zero-crossing rate approach. In *2019 IEEE Conference on Computational Intelligence for Financial Engineering & Economics (CIFEr)* (pp. 1-9). IEEE.

Mahmood, T., Ayyub, R. M., Imran, M., Naeem, S., & Abbas, W. (2020). The behavioral analysis and financial performance of individual investors at Pakistan Stock Exchange. *International Journal of Economics and Financial Issues*, *10*(5), 158.

Majumder, M. M. R., Hossain, M. I., & Hasan, M. K. (2019, February). Indices prediction of Bangladeshi stock by using time series forecasting and performance analysis. In *2019 International Conference on Electrical, Computer and Communication Engineering (ECCE)* (pp. 1-5). IEEE.

Mehtab, S., Sen, J., & Dasgupta, S. (2020, November). Robust analysis of stock price time series using CNN and LSTM-based deep learning models. In *2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA)* (pp. 1481-1486). IEEE.

Mohapatra, P., Raj, A., & Patra, T. K. (2012). Indian stock market prediction using differential evolutionary neural network model. *International Journal of Electronics Communication and Computer Technology (IJECCT)*, *2*(4), 159-166.

Mukherjee, S., & Joshi, S. (2014, May). Author-specific sentiment aggregation for polarity prediction of reviews. In *LREC* (pp. 3092-3099).

Nayak, R. K., Mishra, D., & Rath, A. K. (2015). A Naïve SVM-KNN based stock market trend reversal analysis for Indian benchmark indices. *Applied Soft Computing*, *35*, 670-680.

Nelson, D. M., Pereira, A. C., & De Oliveira, R. A. (2017, May). Stock market's price movement prediction with LSTM neural networks. In *2017 International joint conference on neural networks (IJCNN)* (pp. 1419-1426). Ieee.

Nopp, C., & Hanbury, A. (2015, September). Detecting risks in the banking system by sentiment analysis. In *Proceedings of the 2015 conference on empirical methods in natural language processing* (pp. 591-600).

Oncharoen, P., & Vateekul, P. (2018, August). Deep learning for stock market prediction using event embedding and technical indicators. In *2018 5th international conference on advanced informatics: concept theory and applications (ICAICTA)* (pp. 19-24). IEEE.

Oskooe, S. A. P. (2011, July). The random walk hypothesis in emerging stock market-evidence from nonlinear fourier unit root test. In *Proceedings of the world Congress on Engineering* (Vol. 1, pp. 1-5).

Pant, B., & Bishnoi, T. R. (2001, December). Testing random walk hypothesis for Indian stock market indices. In *Research Paper Presented in UTI Capital Market Conference Proceedings* (pp. 1-15).

Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015). Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. *Expert systems with applications*, *42*(1), 259-268.

Pattewar, T., & Jain, D. (2022, January). Impact of Covid'19 on Indian Stock Prediction by Technical Indicators. In *2022 International Conference for Advancement in Technology (ICONAT)* (pp. 1-6). IEEE.

Picasso, A., Merello, S., Ma, Y., Oneto, L., & Cambria, E. (2019). Technical analysis and sentiment embeddings for market trend prediction. *Expert Systems with Applications*, *135*, 60-70.

Pramod, B. S., & Pm, M. S. (2020). Stock price prediction using LSTM. *Test Engineering and Management*, *83*, 5246-5251.

Qin, J., Tao, Z., Huang, S., & Gupta, G. (2021, March). Stock price forecast based on ARIMA model and BP neural network model. In *2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE)* (pp. 426-430). IEEE.

Qiu, L., & Yang, H. (2020). Transfer entropy calculation for short time sequences with application to stock markets. *Physica A: Statistical Mechanics and Its Applications*, *559*, 125121.

Raji, J. O., Ibrahim, Y., & Ahmad, S. A. (2017). Stock price index and exchange rate nexus in African markets. *International Economic Journal*, *31*(1), 112-134.

Ready, M. J. (2002). Profits from technical trading rules. *Financial Management*, 43-61.

Salles, T., Gonçalves, M., Rodrigues, V., & Rocha, L. (2018). Improving random forests by neighborhood projection for effective text classification. *Information Systems*, *77*, 1-21.

Saxena, A., Bhagat, V. V., & Tamang, A. (2021, August). Stock market trend analysis on Indian financial news headlines with natural language processing. In *2021 Asian Conference on Innovation in Technology (ASIANCON)* (pp. 1-5). IEEE.

Schumaker, R. P., Zhang, Y., Huang, C. N., & Chen, H. (2012). Evaluating sentiment in financial news articles. *Decision Support Systems*, *53*(3), 458-464.

Seetharam, Y. (2013). An analysis of herding behaviour during market cycles in South Africa. *Journal of Economics and Behavioral Studies*, *5*(2), 89-98.

Shahvaroughi Farahani, M., & Razavi Hajiagha, S. H. (2021). Forecasting stock price using integrated artificial neural network and metaheuristic algorithms compared to time series models. *Soft computing*, *25*(13), 8483-8513.

Sharaf, M., Hemdan, E. E. D., El-Sayed, A., & El-Bahnasawy, N. A. (2021). StockPred: a framework for stock Price prediction. *Multimedia Tools and Applications*, *80*, 17923-17954.

Shen, S., Jiang, H., & Zhang, T. (2012). Stock market forecasting using machine learning algorithms. *Department of Electrical Engineering, Stanford University, Stanford, CA*, 1-5.

Shukla, A. K., & Singh, H. (2019). Your OTP is also not secure: Security Issues of Mobile Apps in the E-Commerce. *International Journal of Engineering and Advanced Technology (IJEAT)*, *9*(1).

Silva, T. R., Li, A. W., & Pamplona, E. O. (2020, July). Automated trading system for stock index using LSTM neural networks and risk management. In *2020 international joint conference on neural networks (IJCNN)* (pp. 1-8). IEEE.

Singh, H., & Shukla, A. K. (2021, December). An Analysis of Indian Election Outcomes using Machine Learning. In *2021 10th International Conference on System Modeling & Advancement in Research Trends (SMART)* (pp. 297-306). IEEE.

Singh, H., & Malhotra, M. (2023, March). Artificial Intelligence Based Hybrid Models for Prediction of Stock Prices. In *2023 2nd International Conference for Innovation in Technology (INOCON)* (pp. 1-6). IEEE.

Singh, H., & Malhotra, M. (2023, March). A Time Series Analysis-Based Stock Price Prediction Framework Using Artificial Intelligence. In *International Conference on Artificial Intelligence of Things* (pp. 280-289). Cham: Springer Nature Switzerland.

Singh, H. (2023, April). Prediction of Thyroid Disease using Deep Learning Techniques. In *2023 Second International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT)* (pp. 1-7). IEEE.

Singh, H., & Malhotra, M. (2023). A Novel Approach of Stock Price Direction and Price Prediction Based on Investor's Sentiments. *SN Computer Science*, *4*(6), 823.

Souma, W., Vodenska, I., & Aoyama, H. (2019). Enhanced news sentiment analysis using deep learning methods. *Journal of Computational Social Science*, *2*(1), 33-46.

Srivastava, S., Pant, M., & Gupta, V. (2023). Analysis and prediction of Indian stock market: a machine-learning approach. *International Journal of System Assurance Engineering and Management*, *14*(4), 1567-1585.

Stanković, J., Marković, I., & Stojanović, M. (2015). Investment strategy optimization using technical analysis and predictive modeling in emerging markets. *Procedia Economics and Finance*, *19*, 51-62.

Sullivan, R., Timmermann, A., & White, H. (1999). Data-snooping, technical trading rule performance, and the bootstrap. *The journal of Finance*, *54*(5), 1647-1691.

Sureshkumar, K. K., & Elango, N. M. (2011). An efficient approach to forecast Indian stock market price and their performance analysis. *International Journal of Computer Applications*, *34*(5), 44-49.

Tan, L., Chiang, T. C., Mason, J. R., & Nelling, E. (2008). Herding behavior in Chinese stock markets: An examination of A and B shares. *Pacific-Basin finance journal*, *16*(1-2), 61-77.

Tan, S. D., & Taş, O. (2019). Investor attention and stock returns: Evidence from Borsa Istanbul. *Borsa Istanbul Review*, *19*(2), 106-116.

Ticknor, J. L. (2013). A Bayesian regularized artificial neural network for stock market forecasting. *Expert systems with applications*, *40*(14), 5501-5506.

Vaga, T. (1990). The coherent market hypothesis. *Financial Analysts Journal*, 36-49.

Vargas, M. R., Dos Anjos, C. E., Bichara, G. L., & Evsukoff, A. G. (2018, July). Deep leaming for stock market prediction using technical indicators and financial news articles. In *2018 international joint conference on neural networks (IJCNN)* (pp. 1-8). IEEE.

Wang, J., Sun, T., Liu, B., Cao, Y., & Wang, D. (2018, December). Financial markets prediction with deep learning. In *2018 17th IEEE international conference on machine learning and applications (ICMLA)* (pp. 97-104). IEEE.

Wang, S., Jiang, Z. Q., Li, S. P., & Zhou, W. X. (2015). Testing the performance of technical trading rules in the Chinese markets based on superior predictive test. *Physica A: Statistical Mechanics and its Applications*, *439*, 114-123.

Wang, Z., Hu, Z., Li, F., Ho, S. B., & Cambria, E. (2023). Learning-based stock trending prediction by incorporating technical indicators and social media sentiment. *Cognitive Computation*, *15*(3), 1092-1102.

Waweru, N. M., Munyoki, E., & Uliana, E. (2008). The effects of behavioural factors in investment decision-making: a survey of institutional investors operating at the Nairobi Stock Exchange. *International Journal of business and emerging markets*, *1*(1), 24-41.

Wei, L. Y. (2016). A hybrid ANFIS model based on empirical mode decomposition for stock time series forecasting. *Applied Soft Computing*, *42*, 368-376.

Wichaidit, S., & Kittitornkun, S. (2015, November). Predicting SET50 stock prices using CARIMA (cross correlation ARIMA). In *2015 International Computer Science and Engineering Conference (ICSEC)* (pp. 1-4). IEEE.

Wijesinghe, G. W. R. I., & Rathnayaka, R. M. K. T. (2020, December). Stock Market Price Forecasting using ARIMA vs ANN; A Case study from CSE. In *2020 2nd International Conference on Advancements in Computing (ICAC)* (Vol. 1, pp. 269-274). IEEE.

Williams, J. B. (1938). *The theory of investment value* (Vol. 36). Cambridge, MA: Harvard university press.

Winata, A., Kumara, S., & Suhartono, D. (2021, October). Predicting stock market prices using time series sarima. In *2021 1st International Conference on Computer Science and Artificial Intelligence (ICCSAI)* (Vol. 1, pp. 92-99). IEEE.

Wright Hoffman, G. (1935). GRAHAM, B., and DL DODD. Security Analysis. Pp. xi, 725. New York: McGraw-Hill Book Company, 1934. $4.00. *The ANNALS of the American Academy of Political and Social Science*, *177*(1), 276-277.

Wu, J. M. T., Li, Z., Herencsar, N., Vo, B., & Lin, J. C. W. (2023). A graph-based CNN-LSTM stock price prediction algorithm with leading indicators. *Multimedia Systems*, *29*(3), 1751-1770.

Wu, J. L., Huang, M. T., Yang, C. S., & Liu, K. H. (2021). Sentiment analysis of stock markets using a novel dimensional valence–arousal approach. *Soft Computing*, *25*, 4433-4450.

Xiao, C., Xia, W., & Jiang, J. (2020). Stock price forecast based on combined model of ARI-MA-LS-SVM. *Neural Computing and Applications*, *32*, 5379-5388.

Xiong, L., & Lu, Y. (2017, April). Hybrid ARIMA-BPNN model for time series prediction of the Chinese stock market. In *2017 3rd International conference on information management (ICIM)* (pp. 93-97). IEEE.

Xu, Y., & Keselj, V. (2019, December). Stock prediction using deep learning and sentiment analysis. In *2019 IEEE international conference on big data (big data)* (pp. 5573-5580). IEEE.

Yalçın, K. C. (2010). Market rationality: Efficient market hypothesis versus market anomalies. *European Journal of Economic and Political Studies*, *3*(2), 23-38.

Yang, J., Wang, Y., & Li, X. (2022). Prediction of stock price direction using the LASSO-LSTM model combines technical indicators and financial sentiment analysis. *PeerJ Computer Science*, *8*, e1148.

Yazdi, S. H. M., & Lashkari, Z. H. (2013). Technical analysis of Forex by MACD Indicator. *International Journal of Humanities and Management Sciences (IJHMS)*, *1*(2), 159-165.

Yıldırım, D. C., Toroslu, I. H., & Fiore, U. (2021). Forecasting directional movement of Forex data using LSTM with technical and macroeconomic indicators. *Financial Innovation*, *7*, 1-36.

Zhang, G., Xu, L., & Xue, Y. (2017). Model and forecast stock market behavior integrating investor sentiment analysis and transaction data. *Cluster Computing*, *20*, 789-803.

Zhang, L., Ding, X., Hou, R., & Tao, Y. (2019, November). Integrating Harmony Search Algorithm and Deep Belief Network for Stock Price Prediction Model. In *2019 6th International Conference on Systems and Informatics (ICSAI)* (pp. 631-636). IEEE.

Zhou, Z., Zhao, J., & Xu, K. (2016). Can online emotions predict the stock market in China?. In *Web Information Systems Engineering–WISE 2016: 17th International Conference, Shanghai, China, November 8-10, 2016, Proceedings, Part I 17* (pp. 328-342). Springer international publishing.