

CHAPTER II

REVIEW OF LITERATURE

2.1 Introduction

Stock prediction is the process of trying to regulate the future value of stock values. The effective prediction of the stock price in the future is a challenging task that makes more profit to the customers and investors. The stock prediction is performed with the help of past historical datasets. The ANNs are a very commanding method to predict the stock market's future behavior due to its higher problem-solving ability.

On the other hand, the stock data is affected by numerous characteristics such as nonlinearity, noise, redundancy, etc. In this circumstance, constructing the perfect prediction methods is a challenging task. Hence, frequent pattern mining is a task of discovering relationships or correlated features among the given datasets. The frequent pattern mining method is incorporated with ANN to enhance the prediction accuracy with correlated features. In this chapter, we concentrate brief review regarding stock price prediction using ANN and frequent patterning mining.

Stock prediction plays a significant role in financial stock market research. If a stock market can be predicted efficiently, the customer or investors can achieve more returns. However, there are many aspects affecting the outcome of the prediction and there are no proper universal techniques. Therefore, a variety of analysis techniques have been used by numerous academics to predict stock trends.

2.2 Related work on ELM-based prediction methods

The following subsections discuss the many research articles that refocused on ANNs-based prediction methods. D. Mohanty et al. (2021) [22] developed an amalgamated hybrid model incorporating autoencoder (AE) and kernel KELM is suggested to enhance the caliber of financial market forecasting even more. This study primarily focuses on accurate financial market prediction; stock price prediction serves as a means of reducing the significant losses that investors incur and of analyzing profitability via the use of buy and sell amounts. The suggested method's main benefit over the traditional SAE is its reliable and error-reduced forecast of various financial markets.

S. Deng et al. (2014) [23] propose an advanced prediction approach by integrating the Boosting, Bagging, and NSGA-II methods for the Shanghai Stock Exchange index and the Shenzhen Component index. Our goal is to examine the efficacy of stock prediction using multiple stockholder sentiment features. The proposed strategy performs better than the benchmark methods in terms of hit fraction, accumulated arrival, and extreme drawdown, according to the experimental data. It shows that in price prediction, the suggested strategy can achieve high accuracy, low danger, and constant revenue. Furthermore, the suggested model's interpretability was enhanced by the SHAP technique, which can serve as a useful guide for market participants to identify significant sentiment components and make appropriate judgments.

X. Li et al. (2016) [24] describe the architecture and design of the transaction signal stage based on ELM to forecast based on those two data sources simultaneously. Extensive experimental comparisons between ELM and cutting-edge learning algorithms, such as SVM and BPNN, have been conducted using contemporaneous news archives and intra-day tick-by-tick data of the H-share market. The findings indicate that:

- 1) RBF-ELM and RBF-SVM reach comparable levels of accuracy
- 2) RBF ELM outperforms RBF SVM in terms of prediction swiftness. With the signals, simulations of an initial trading strategy are run. Findings indicate that a strategy with more precise signals will yield more with low-risk

S. Das et al. (2022) [25] suggested to use Cluster-based Modified Crow Search Algorithm (CMCSA), Optimized ELM (OELM), and Intrinsic Time-Scale Decomposition (ITD) to predict random nonlinear stock market price. To separate the data into a residual component and a few Proper-Rotation-Components (PRCs), ITD is used. Non-stationary stock price data are transformed into stationary data using ITD, which makes the data easier to analyze and more consistent.

The Crow Search Algorithm (CSA) is modified to pick the ideal weight and ELM biases more effectively, resulting in the proposal of the CMCSA. Subsequently, the PRCs and residual components are predicted separately using the optimized ELM, and these predictions are joint to forecast the closing price of short-term stocks. Solving benchmark problems tests and validates the effectiveness of the CMCSA. According to the experimental investigation, the suggested model performs better.

D. Wu et al. (2021) [26] suggested using DWT and ELM-based denoising in tandem to forecast stock trends. The suggested approach was applied to forecast the 400-stock market's trend in China. The suggested method's prediction results show exceptional performance when compared to 12 ML methods, providing solid evidence of the effectiveness of DWT-based denoising for stock movements.

A. Y. Fathi et al. (2021) [27] suggest a hybrid model to estimate daily closing prices by combining the BPNN with the single-spectrum analysis (SSA). Using the SSA, the model first breaks down the stock prices. Afterward, BPNNs are trained with the extracted components to predict future prices. The hybrid SSA-BPNN model outperforms the BPNN in terms of prediction, proving the SSA's capacity to uncover hidden data and lessen the noise impact on the initial time series.

Y. Wu et al. (2021) [28] developed a model that can forecast multi-day stock prices by optimizing BPNN using the GA. The GA-BP model is then enhanced with the principal component analysis (PCA) algorithm, which aims to offer a useful method for assessing the market risks associated with the A+H shares. According to the experimental data, the technique has the best forecast result on the price of BC (Back of China) for A shares; the average prediction errors of the starting price, extreme price, least price, and closing price are, respectively, 0.0236, 0.0262, 0.0294, and 0.0339. The developed model has the greatest impact on China Merchants Bank's (CMB) price prediction for H shares

S. Das et al. (2022) [29] suggested combining the ELM with a modified crow search algorithm (CSA) for stock forecasting efficacy. By resolving 12 benchmark issues, the PSO-based group-orientated CSA (PGCSA), a modified CSA that has been suggested, is shown to perform better than other existing algorithms. To increase the efficacy of traditional ELM, the PGCSA method is utilized to obtain the appropriate weights and biases of ELM.

Using performance metrics, technical indicators, and a hypothesis test (paired t-test), the hybrid PGCSA ELM model's ability to predict the closing price of seven distinct stock indices the next day is observed. Data from the COVID-19 outbreak is taken into consideration while analyzing the seven stock indices. This model is evaluated by contrasting it with current methods that have been put forth in published papers. According to the simulation findings, the PGCSA-ELM is a good tool for predicting the closing price of the following day.

F. Zhang (2021) [30] built an ELM hybrid technique to analyze performance and predict stock price using the ELM, a recent modification of the classic FFNN with a single hidden layer. Its process speed is greatly better based on a strong mathematical background and its ability to avoid problems like local minimum. The viability of using the ELM model for stock market analysis is tested by comparing the expected output with actual data. The outcome shows that the ELM model is useful for forecasting the trend of stock price variation, but there is some error in the peak value prediction, which may call for more model adjustment. Overall, using the ML model ELM to forecast stock price and produce desired results, both broadens the field deployment of the ELM and provides a fresh perspective on stock market performance.

L. Xiong et al. (2017) [31] created a novel hybrid ARIMA-BPNN using technical indicators to forecast four individual stocks in the software and information services sector that comprise both the main board market and the growing enterprise market. The results of the experiment demonstrate that the suggested method outperforms other methods in terms of one-step-ahead forecasting accuracy, with values for each series of 78.79%, 72.73%, 59.09%, and 66.67%, respectively.

M. Vochozka et al. (2020) [32] seek to employ neural networks to improve prediction management when anticipating the evolution of share prices over time. The information on CEZ, a.s. Share prices for the contribution came from the database of the Prague Stock Exchange. Data on stock prices are accessible from 2012 to 2017. Based on the statistical analysis of the data, the networks with the best features are kept, and all are useful in real-world scenarios. MLP networks outperform SVR and BPNN networks in terms of stable performance across all data sets. In terms of the overall evaluation, the greatest significant difference between the maximal and lowest predictions is where the deviation of 2.26% is found. In terms of one stock's price, it isn't always important. However, the difference could appear substantial if you're buying or selling a lot of stocks. In light of this, it is advised to apply ML P 1-2-1 and ML P 1-5-1 in practice. A minimal, pessimistic prediction is always represented by the first network. Of the suggested networks, the second one is maximally optimistic. The interval of the difference between the optimistic and pessimistic projections should match the actual scenario.

H. Hu et al. (2018) [33] proposed an enhanced sine cosine algorithm (ISCA) to optimize the weights and basis of BPNN. The enhanced SCA (ISCA) adds a parameter to the SCA. Consequently, a new network called ISCA-BPNN is created by combining ISCA and

BPNN to forecast the directions of the opening stock prices for the DJIAI and the S&P 500 Index, respectively. To improve stock prediction, Google Trends data is also taken into account. They examined two different kinds of predictions: Type I, which was made without the use of Google Trends, and Type II, which was made with the help of Google Trends. The hybrid ISCA-BPNN model is used to confirm that the direction of stock prices may be predicted.

According to the experimental data, ISCA-BPNN performs much better for Type II and better than other methods in terms of forecasting the direction of the opening price for both types. The S&P 500 Index and the DJIAI have hit ratios of 86.81% and 88.98%, respectively, for ISCA-BPNN using Google Trends. The findings indicate that one can use Google Trends to forecast the direction of the stock market index.

D. Jia et al. (2022) [34] examine the relationship between internal governance and earnings management using research samples of China's listed companies on GEM from 2011 to 2020. It does this by introducing the BPNN and GBDT (Gradient Boosting Decision Tree) models, which are then compared to the empirical findings of the traditional multiple linear regression to determine their validity and predictive power in the field of earnings management research. The findings indicate what follows.

- 1) There is a significant rate of experimental data distortion in the multiple linear regression model's matching impact while analyzing GEM. Nonetheless, compared to the multiple linear regression model, the gradient lifting tree model and BPNN have far superior predictive power.
- 2) The study of accrual earnings management is more suited for the gradient lifting tree model, but the study of actual earnings management is better suited for the BP neural network.

M. Jiang et al. (2022) [35] suggested a new ensemble model for stock price prediction that combines the improved harmony search (IHS) algorithm, ELM, empirical mode decomposition (EMD), and variational mode decomposition (VMD), VMD-ELM-IHS). These models are called EMD-ELM-IHS and VMD-ELM-IHS, respectively. Additionally, to illustrate the effectiveness and performance of the suggested models, the outcomes were contrasted with those derived from alternative techniques. The findings demonstrate that, in comparison to the other models, the suggested models perform better in terms of accuracy

and stability. Additionally, they discovered that the sliding window and training set sizes significantly affect the performance.

S. Mohan et al. (2019) [36] increased the precision of stock price forecasts by compiling a substantial amount of time series data and applying DL models to analyze it in connection with relevant news stories. They have collected daily stock prices for S&P 500 businesses going back five years, in addition to over 265,000 financial news stories about these companies. Because of the magnitude of the dataset, cloud computing proved to be a priceless tool for them in training prediction models and making real-time inferences for a certain stock.

2.3 Related work on prediction models based on sentiment analysis

N. Jing et al. (2021) [37] suggested a hybrid prediction model that blends a sentiment analysis model and a DL method. The author used CNN to classify the hidden attitudes of investors that were taken from a significant stock forum. then use the LSTM approach to analyze the stock market's technical indicators and the sentiment analysis findings from the previous stage to propose a hybrid model.

To confirm the efficacy and suitability of the suggested methodology, it also carried out real-world trials on the Shanghai Stock Exchange (SSE) at three-month intervals across six important industries. According to the experiment results, the hybrid strategy outperforms the single techniques without sentiment analysis in terms of stock price prediction. The recommended model has also outperformed the baseline classifiers in terms of investor sentiment classification.

Z. Jin et al. (2020) [38] presented a DL-based stock prediction scheme that takes emotional inclinations of investors into account. To increase the accuracy of the model's stock forecast, the authors first suggested taking investors' emotions into account. Second, it is exceedingly difficult to make an accurate prediction since the stock pricing sequence is a complex temporal sequence with varying scales of fluctuations. To improve prediction accuracy, the authors suggest using EMD to progressively break down the intricate sequence of stock prices.

Third, they use LSTM because of its memory function's ability to analyze relationships between time series data. They made additional revisions by implementing an attention mechanism to concentrate more on the most important facts. The results of the

testing validate that the updated LSTM model can decrease time delay in addition to increasing prediction accuracy. It has been established that investors' emotional tendencies can effectively enhance expected outcomes; the addition of EMD can enhance inventory sequence predictability; and the attention mechanism can assist LSTM in effectively extracting relevant information and current mission objectives from the information ocean.

R. Gupta et al. (2020) [39] examined the effect on stock price prediction of sentiment expressed on StockTwits. A relatively new microblogging platform called StockTwits is gaining popularity as a place where people can express their opinions and conversations about stocks and financial markets. In particular, the scientists used a set of text features and machine learning techniques to assess the contents of StockTwits tweets and extract financial sentiment. Next, the relationship between the daily sentiment aggregate and the daily movement of the stock price is examined. Lastly, to increase the precision of the stock price movement forecast, sentiment data is added to historical stock time series data. Experiments on five firms using daily stock data and nine months of StockTwits data illustrate the efficacy of the proposed work.

S. Wu et al. (2022) [40] suggested S_I_LSTM, a stock price prediction technique that takes into account the mood of investors as well as several data sources. They start by preprocessing the various data sources they find online by crawling them. Technical indicators, non-traditional data sources like stock posts and financial news, and historical data. Next, apply the CNN-based sentiment analysis technique for non-traditional data to determine the investors' sentiment index. Lastly, using an LSTM and combining sentiment index, technical indicators, and historical stock transaction data as the feature set of stock price prediction, forecast the China Shanghai A-share market. According to the experiments, the mean absolute error can reach 2.386835, which is healthier than old-style approaches, and the predicted closing price is closer to the genuine closing price than the single data source. They used the actual data sets of five listed businesses to confirm the efficacy.

P. Mehta et al. (2021) [41] created and put into use a forecast accuracy method that took into account public opinion in addition to other factors. To predict future stock values, the suggested algorithm takes into account news, opinions, public mood, and past stock prices. ML and DL techniques, such as SVM, MNB classifier, LR, NB, and LSTM, were used to conduct the tests. The outcomes confirmed that the suggested methodology was successful.

Y. Mehta et al. (2021) [42] developed a model with the least error is the best and most recommended technique for prediction. Three different models were employed for this study, and sentiment analysis was also done on the tweets about the company or the stock. The classification's outcomes have provided investors with a fresh perspective on where to place their bets and a clear, intelligent understanding of the market's erratic ups and downs. For each stock, the ARIMA model provides the best accuracy.

N. Darapaneni et al. (2022) [43] presented to forecast share price movement in the future by utilizing past prices and sentiment data that were available. The exercise involved the employment of two models, the first of which was the LSTM model, which used historical prices as the independent variable. Major parameters for the Random Forest Model employed in the second portion included gold prices, oil prices, the exchange rate between the US dollar and the Indian rupee, and sentiment analysis obtained with the Intensity Analyzer. Additionally, securities yields were incorporated into the algorithm to increase its accuracy. As the final result, the prices of four stocks—Reliance, HDFC Bank, TCS, and SBI—were forecasted through the use of the two models mentioned above. The RMSE metric was used to assess the outcomes.

K. Srijiranon et al. (2022) [44] To anticipate the closing price of the Thai stock market one step ahead of time, a hybrid prediction model known as PCA-EMD-LSTM was developed by combining PCA, EMD, and LSTM. In this study, news sentiment analysis was also used to enhance the functionality of the FinBERT-based suggested framework, which is based on financial and commercial news. The study scrutinized stock market pricing experiments conducted in Thailand between 2018 and 2022, employing a range of statistical indicators as evaluation standards. The findings collected demonstrated that, when compared to baseline approaches for stock market price prediction, the suggested framework produced the best results. Moreover, using news sentiment analysis can improve how well the initial LSTM model performed.

R. Batra et al. (2018) [45] utilizing sentiment analysis's advantages for the stock market, they have examined tweets on Apple items that were taken from the social media platform StockTwits between 2010 and 2017. Tweet sentiment is analyzed using Support Vector Machines (SVM) to get a tweet's sentiment score. Each tweet is therefore classified as either bullish or bearish. Next, an SVM model is constructed using sentiment score and market data to forecast the movement of stocks the following day. The results indicate a

positive relationship between market data and people's opinions, and the suggested work has a 76.65% accuracy rate.

I. K. Nti et al. (2020) [46] developed new forecast techniques based on the forthcoming stock value for a time window was projected using a predictive framework that was tested using stock data that was received from the Ghana Stock Exchange (GSE) between January 2010 and September 2019. Based on Google trends, Twitter, forum posts, web news, and a merged dataset, the authors found that the accuracy ranges were (49.4–52.95%), (55.5–60.05%), (41.52–41.77%), and (70.66–77.12%). As a result, they noticed an improvement in the prediction model's accuracy when multiple stock-related data sources were integrated. Also a strong correlation between social networking sites and stock market behavior. In light of the study's findings, they consequently suggested that stock market investors make use of data from online financial news, tweets, forum discussions, and Google trends to accurately predict future stock price movement and create successful portfolio and investing strategies.

Y. Guo (2020) [47] developed a new LSTM that uses sentiment analysis to predict stock prices. The author gets news stories from the New York Times, applies sentiment analysis to the title and body of the story, and then uses an LSTM to integrate quantitative sentiment score with stock historical stock basic attributes to anticipate future stock closure price and stock return. The primary goal is to compare the forecast outcome of a model that takes sentiment elements into account with one that just takes historical stock fundamentals into account. The LSTM model for three significant US corporations serves as the basis for this experiment, which demonstrates the effectiveness of LSTM neural networks in long-term prediction. The final results verify that the model that takes into account the sentiment influence from websites has a higher prediction accuracy.

A. Derakhshan et al. (2019) [48] tackled the issue of model-based opinion mining by presenting a part-of-speech model to extract user opinions and assess it in two distinct datasets in Persian and English, with the Persian dataset obtained from an Iranian stock market social network for this paper. This model outperformed other approaches that use explicit sentiment labeling for comments in terms of accuracy when predicting the stock market.

A. Pathak et al. (2019) [49] seek to integrate several current methods into a significantly more reliable prediction model that can handle a variety of situations when

investing may be advantageous. Existing methods, such as neural networks or sentiment analysis techniques, may be overly limited in their thinking and produce inaccurate results in a variety of situations. This prediction model can generate recommendations that are more precise and adaptable by combining the two methods. Technical indicators integrated into the investment will help the investor reduce risk and maximize returns.

M. Kim et al. (2018) [50] aim to develop a model by simulating the methods used by analysts, investors, and traders to assess stock investing strategies. To comprehend and forecast market behavior for Google stock, a combination of fundamental analysis utilizing news headlines and technical analysis utilizing readily available numerical data about equities is used. Sentiment analysis is utilized in this context to interpret news reports about the stock in addition to the time series data that is currently available as input for an LSTM neural network. It has been noted that using this method produces a more generalized, accurate, and intuitive model that can be used to anticipate stock market movements.

S. Albahli et al. (2022) [51] use a unique StockSentiWordNet (SSWN) model to analyze Twitter postings and Google Finance data to predict the future behavior of the stock markets, one of the major financial markets, in a specific period, such as hourly, daily, weekly, etc. To train ELM and RNN for stock price prediction, the proposed SSWN model expands the SentiWordNet (SWN) conventional opinion lexicon with phrases that are uniquely connected to the financial markets. The Sentiment140 and Twitter datasets were used for the studies, which yielded an accuracy value of 86.06%. The work performs better overall in terms of accuracy than state-of-the-art methods, according to the findings.

R. Kumar et al. (2021) [52] committed to creating a radial basis function network (RBFN)-based intelligent prediction model. Back-propagation (BP) is a technique used in the development of learning algorithms to fine-tune their parameters. The findings of the simulation study show that the performance of RBFN is superior to that of the ML FFNN model. The performance of the suggested technique is also compared with that of the multi-layered feed-forward neural network (ML FFNN), which has just one hidden layer.

K. A. Rashedi et al. (2021) [53] suggested a mathematical approach to identify anomalies in the Saudi stock market's daily return using artificial intelligence (Tadawul). A data point that differs excessively from the remaining observations in a data sample is called an outlier. The inflation rate, repo rate, and oil prices were chosen as input variables

according to the Engle and Granger Causality test. The authors initially employed the Tukey approach to identify outliers in the Tadawul stock return data that was gathered between October 2011 and December 2019 before constructing the mathematical model.

The stock return data was divided into two groups by the authors: nonoutliers and outliers. These data are also utilized in conjunction with the particle swarm optimization algorithm to train ANNs. The MSE was used to evaluate the suggested model's performance. The suggested model is shown by a 0.05 mean squared error value. Direct detection of outlier values from the repo rate, inflation rate, and oil prices is possible using the suggested model. When creating and implementing intelligent optimization approaches to address issues in business processes, the suggested model can be useful.

V. Saravanan et al. (2023) [54] employed LSTM, a time series forecasting technique, to assess the model's performance and anticipate the future stock market values of Facebook (FB) and Twitter (TWTR) on the NYSE. Predicting changes in stock market values also heavily relies on sentiment analysis, which entails interpreting data from sentiment-rich sources like news and social media. Modern data is so abundant that human data evaluation is impractical, calling for the use of artificial intelligence (AI) and data mining tools. This study also looked into how emotion classification affects stock market prices. Rather than depending on human feature engineering, the DL technique used an LSTM to learn the features. This study's NLP model blends real-time sentiment analysis visualization with LSTM, an interesting subfield of ML that leverages historical data to forecast future values.

C. Manjunath et al. (2023) [55] created a hybrid forecasting model using PCA-ML (ML). The experimental results indicate that the technical aspects are trend signals and that the PCA technique with ML models performs better in prediction than the comparable models. Experiments on the Nifty 50 index using the first three principal components demonstrate that the RF model achieves this.

P. R. Mohapatra et al. (2023) [56] By combining an improved Online Sequential OSELM with Multiple Criteria Decision-Making (MCDM), a novel stock index trend forecasting model is created. The two goals of the suggested model are to forecast future stock index prices and examine any upward or downward trends in these price estimates. Appropriate activation function selection for OSELM is treated as an MCDM problem since the activation functions utilized in it have a significant impact on its performance. OSELM's training phase incorporates a hybrid CSA (hCSA) to further optimize the system's

performance. The suggested hCSA can strike the ideal balance between exploration and exploitation, enhancing convergence, by adding the chaotic map, mutation operator, and catfish behavior to the original CSA's search process and position update method.

The suggested trend predictor model is tested empirically using pre-COVID and COVID-era historical data from three stock indices, including the DJIA, S&P 500, and BSE SENSEX. The hCSA-OSELM model performs better in the majority of test cases than the most advanced baseline models across all evaluation criteria. The proposed model can achieve MSE improvements of 4 to 6%, 25 to 31%, and accuracy increases of 0.4 to 0.8%, 0.9 to 1.3%, during the pre-COVID and COVID time-frames, respectively, when compared to the second-best baseline model. The enhanced performance of the suggested model is further demonstrated by the statistical test. The suggested model can be employed for financial time-series forecasting amid daily volatility as well as extremely volatile markets, as demonstrated by the strong and dependable MCDM-based model selection, superior prediction, and classification results.

C. M. Liapis et al. (2023) [57] Modern techniques for financial time series forecasting with sentiment analysis are compared in this article. 67 distinct feature setups, comprising sentiment scores and stock closing prices, were investigated using a wide range of metrics and datasets in a comprehensive experimental procedure. Throughout two case studies—one comparing input feature sets, the other comparing methods—a total of thirty cutting-edge algorithmic techniques were employed. The combined findings show the following: (1) a conditional improvement in model efficiency following the addition of sentiment sets in specific prediction time frames; (2) the ubiquity of the suggested strategy. The statistical test provides more evidence of the improved performance of the proposed model. The robust and consistent MCDM-based model selection, excellent prediction, and classification results show that the proposed model may be used for financial time-series forecasting in the middle of daily volatility as well as exceptionally volatile markets.

A. Rodrigues et al. (2023) [58] propose a hybrid approach that incorporates sentiment analysis with ensemble learning techniques to predict stock market prices. The proposed approach consists of four main steps:

- 1) Sentiment analysis of news and social media data related to a particular stock
- 2) Fetching historical stock data for the said stock;
- 3) Feature extraction using various technical indicators;
- 4) Ensemble learning using a combination of multiple ML models.

The proposed approach was evaluated based on the stock prices of five key companies in the technology industry. The results showed that the hybrid approach outperformed individual ML models and traditional time-series forecasting methods in terms of accuracy and consistency. The ensemble learning technique aided in reducing the effect of overfitting and increasing the robustness of the model. The sentiment analysis component contributed to enhancing the prediction accuracy by providing insights into the market's sentiment towards a particular stock. Ultimately, the research project demonstrates the potential of using sentiment analysis and ensemble learning techniques to predict stock market prices. The proposed approach can be used by financial analysts and investors to make informed decisions and mitigate the risks associated with stock investments.

P. Jishtu et al. (2023) [59] provide new, reliable benchmarks in the extremely erratic stock market. However, there are numerous resources available to learn about this dynamic shift and get ready for it as technology advances. This essay focuses on many techniques for rapidly recognizing market trends. Because it incorporates pre-processing of the stock market dataset, various feature engineering techniques, and the integration of a bespoke DL-based system for stock market price pattern forecasting, the proposed strategy is considered comprehensive.

The model with the lowest level of error is the most effective and recommended approach for prediction. They employed three different models and performed sentiment analysis on news stories that mentioned the company or the stock to carry out this investigation. The classification's outcomes have provided investors with more data to aid in their decision-making over where to place their bets as well as a lucid and insightful understanding of the market's erratic ups and downs.

J.-Y. Huang et al. (2023) [60] suggest a way to forecast changes in stock prices by combining sentiment from social media, GA, and DL. To find chip-based indicators that are strongly associated with changes in stock prices, it first uses a hybrid GA (HGA) in conjunction with ML. These indicators are then fed into an LSTM to create a prediction model. Subsequently, this research suggests five sentiment measures to examine the impact of PTT social media on TSMC's stock price. A grey relational analysis (GRA) is then conducted to determine which sentiment factors are more strongly associated with changes in stock price. The chosen chip-based indicators and the sentiment variables are then integrated as input to create the LSTM prediction model. This work uses the Taguchi technique to optimize the hyper-parameters to increase the effectiveness of the LSTM analysis. The

findings demonstrate that the suggested strategy of establishing an LSTM prediction model with input from social media sentiment variables and HGA-screened chip-based variables can significantly increase prediction accuracy.

N. Zaman *et al.* (2023) [61] suggested a machine learning model that, to improve prediction accuracy, combines historical data with outside variables including social media attitudes, trends in gold and oil, and financial news data. To verify the efficacy of the suggested model, this study examined the effects of COVID-19 on businesses using datasets from the stock markets of HPQ, IBM, ORCL, and MSFT. Using oil and sentiment datasets, these experimental findings demonstrated the greatest accuracy of 87.2%. They also found that social media has a big impact on IBM stock prices and that the GBM (Gradient Boosting Classifier) classifier consistently yielded reliable findings.

P. K. Mallick *et al.* (2023) [62] An innovative hybrid multi-layer decomposition method is created. This work proposes a decomposition method that combines the separate techniques of EMD and VMD. This study adds to the body of knowledge by using the VMD to further decompose the higher frequency signals that are derived from the EMD-based decomposed signal. The twofold decomposition technique increases prediction accuracy, according to the result analysis. In the sphere of stock market prediction, this is a novel introduction. The suggested model's prediction accuracy is verified through its application to three distinct sets of stock market data. To anticipate the closing price one day in advance, historical data (close price) is used.

Along with the suggested strategy, a comparative analysis of various previously used techniques including BPNN, SVM, ANN, and ELM is done. The kernel factors are optimized with the application of the Fire Y technique. It is noted that the suggested hybrid model performed better than the other approaches mentioned.

P. Rajendiran *et al.* (2023) [63] One technique to identify stock market activity is sentiment analysis, which has been developed. Several components are experiencing fluctuations as a result of the unpredictability of the stock market. A non-linear relationship and several factors might influence the trend of a sequence.

2.4 Summary

The prediction of the stock market is a topic that is appealing to academics and practitioners alike. By compiling client reviews and categorizing them as either good or

negative, the prediction is executed. Sentiment analysis was used to introduce several statistical and econometric methodologies for stock market prediction. Sentiment analysis tools suffer from errorless stock market predictions. A direct reason for lower categorization accuracy is found in stock market indicators.

Early on, statistical techniques like moving average (MV), ARIMA, GARCH, and others are widely used for stock market prediction. However, these models do not capture the impact of sentiment and perform poorly when dealing with enormous datasets. SVM and BPNN were used to address the shortcomings of the statistical approach because of their high classification rate and function approximations. These methods, however, require a lot of fine-tuning, are computationally costly for large datasets, and have trouble with noisy sentiment data. However, deep learning techniques such as LSTM and others take market psychology into account and reflect long-term dependencies in-stock patterns. These models require a lot of preprocessing to remove sentiment noise, which raises the model's complexity and computing cost.

Additionally, a new optimized prediction based on ELM with DWM and sentiment analysis is developed in the current research study. With high-dimensional data, including sentiment features, the suggested approach scales well and is better at capturing non-linear correlations. To prevent local minima and lessen overfitting, it trains in a single step. Discovered a method that has a quicker rate of generalization than LSTM and requires less training data. Although ELM + sentiment analysis has a much lower computational cost and a faster inference speed, it nevertheless achieves a similar sentiment integration.