Indonesian Journal of Electrical Engineering and Computer Science
Vol. 39, No. 1, July 2025, pp. 575~584
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v39.i1.pp575-584

Predictive modeling for equity trading using sentiment analysis

Chetan Gondaliya[1], Abhishek Parikh[2]
[1]Swarrnim School of Computing and IT, Swarrnim Startup and Innovation University, Gandhinagar, India
[2]School of Liberal Arts and Management, P P Savani University, Kosamba, India

Article Info

Article history:

Received Apr 23, 2024
Revised Feb 24, 2025
Accepted Mar 26, 2025

Keywords:

Equity returns
Natural language processing
Sentiment analysis
Technical analysis
Trading

ABSTRACT

Warren Buffett's investment philosophy highlights the importance of generating wealth through available capital, but investors require more advanced tools for informed decision-making. Current research is focused on developing a modeling technique that leverages computer algorithms, including sentiment analysis. This method evaluates public sentiment about companies through social media, aiding investors in identifying promising stocks and safeguarding their wealth against unfavorable market conditions. In India, the banking, real estate, and pharmaceutical sectors are among the most robust and rapidly growing industries; however, deciding to invest in these sectors remains debatable. To address this, the proposed study aims to develop a hybrid prediction model that combines sentiment and technical analysis to uncover short-term trading opportunities. This model utilizes a two-layer ensemble stacking technique, training three distinct machine-learning algorithms in the first layer and aggregating their outputs in the second layer. The proposed model significantly outperforms traditional methods in terms of accuracy, enabling investors to make more confident and profitable decisions in the Indian stock market.

Corresponding Author:

Abhishek Parikh
School of Liberal Arts and Management, P P Savani University
394125 Dhamdod, Kosamba, Surat, Gujarat, India
Email: f13abhishekp@iima.ac.in

1. INTRODUCTION

The financial market has grown increasingly competitive, with online stock trading becoming crucial for the economy. However, analyzing the vast amounts of data available to investors takes time and effort. The stock market is influenced by many factors, making stock market prediction a difficult task. The proposed research aims to identify different possibilities in short-term stock market prediction using supervised machine learning (ML) algorithms. Predicting stock prices for a short-term duration using various social media sentiment sources is an ongoing work [1], [2].

Investors generate a large volume of stock market data, which is continuously increasing with the advent of the latest technologies. Processing this vast volume of unstructured textual data is difficult for humans. Therefore, computational approaches using ML algorithms are used to process the data and develop algorithms and scripts based on statistical techniques. ML algorithms have better capabilities than classical approaches in finding hidden features through self-learning [2].

The stock market is non-linear, unpredictable, dynamic, and influenced by various external factors. Computational intelligence approaches can be used to design models using classical or ML techniques, enhancing the prediction power of the proposed model for better trading decision-making. Investors worldwide are interested in investing in the stock market for quick fund growth. Therefore, computational intelligence approaches can help investors make better trading decisions by interpreting various factors that impact the stock market movement [3]-[5].

Stock market behavior is influenced by various factors, including the country's economic situation, credit policies, company fundamentals, domestic and foreign news, social sentiment, technical aspects, demand and supply, and more [3]-[5]. Analyzing this vast amount of structured and unstructured data in a short amount of time is a difficult task for individual investors.

The primary objective of this research is to improve the performance of stock market forecasting techniques for short-term periods. The proposed study aims to design a model to predict future stock prices based on technical and sentimental factors. The study focuses on the Indian stock market only, and the data collection process involves two types of data: sentiment data and technical data. Sentiment data is obtained from various sources, such as forum discussions, RSS feeds, Twitter tweets, and news portals. Technical data is directly collected from Yahoo Finance.

2. LITERATURE REVIEW

Predicting stock prices can be challenging due to factors like the economy, investor sentiment, and political developments. The efficient market hypothesis claims that predicting stock prices is impossible, but there is always the presence of inefficiency that helps to predict stock prices to get abnormal returns [6]-[8].

Statistical techniques are commonly used for stock price prediction. The most widely used techniques include exponential smoothing, autoregressive integrated moving average (ARIMA), and regression method. However, ML algorithms have significantly improved the accuracy of stock market price predictions [2], [9], [10]. Support vector regression (SVM) is a popular machine-learning technique for predicting stock prices. SVM is a widely used ML technique to predict stock prices [11], [12]. One of the studies recommended that future research include more technical analysis indicators and a larger gamut of stocks and markets [2]. Other research used the same approach to predict stock price movements [13]. The model proposed by the study used the multiple regression method on 24 economic indicators.

Some of the researchers had presented a novel dataset that combines technical stock market data with news sentiment and key economic indicators such as inflation, GDP, exchange rates, and interest rates or integrates technical indicators, contextual information, and financial data, employing a heuristic stock selection algorithm to find stocks with high predicted daily returns. Their evaluation across stock and cryptocurrency markets shows enhanced performance compared to existing methods. They emphasize the significant role of news sentiment on stock prices and highlight the need for further research to address the complexities of non-linear and non-stationary data in stock market predictions [14], [15].

A review of articles on ML in stock price forecasting found that many studies favored shorter time frames. Random forest (RF), SVM, artificial neural network (ANN), decision tree (DT), logistics regression (LR), and K-nearest neighbor (KNN) were the most commonly used algorithms [4], [16], [17].

Research gap identification and proposed model: Researchers have used ML algorithms like SVM and RF to forecast stock prices. Researchers also used sentiment analysis based on social media or news feeds to identify expressed sentiment toward specific stocks for forecasting. The study by Maini and Govinda [18] found that RF performed better than SVM. Sentiment analysis can also be used to suggest improvements for products and services [19]. Naïve Bayes classifier can help to classify pessimistic and optimistic sets based on the underlying sentiment for stock price prediction [20].

A comprehensive literature review of sentiment analysis and ML algorithms has provided insight and identified research gaps [21]-[23]. The review found that most international studies did not include the Indian market as sample data.

The geographical inclusion of the world's second-largest market will create a significant gap, as India is among the top ten largest stock markets globally [24]-[26]. Additionally, few studies covered 2020, which was affected by the COVID-19 pandemic [27]-[30]. The review also identified a need for more diverse data sources, including social media platforms and technical indicators after any crisis period, such as COVID-19 [5], [9], [10], [31].

Another study selected ten reviews, including over 379 studies, based on screening 69 titles. They emphasize the use of SVM, long short-term memory (LSTM), and neural networks (NN) [32]. One can also use ANN, SVM, and LSTM networks to predict stock market trends [33]. Historical stock prices were the primary data source, with accuracy being the key performance metric. The current study also focused on historical prices for prediction modeling and accuracy as a comparison matrix to emphasize the importance of the hybrid model in stock market prediction [32], [33].

Various researchers have shown that SVM, RF, and LR provide better accuracy during COVID-19 and other crises, aiding in creating early warning signals to prevent wealth destruction. They emphasize the importance of hybrid models over single prediction methodologies to benefit policymakers, investors, and corporations [34]-[36].

In the other findings, it was evident that increasing the use of hybrid approaches that combine ML and statistical methods enhances prediction accuracy. Authors recommend adopting PRISMA 2020 so researchers, editors, and reviewers can achieve more accurate reporting in their work [37]-[40]. Hence, the current study does not include only statistics but text mining as a tool in a proposed hybrid model.

Review research paper with review of 57 research articles on predicting stock market behavior using time series analysis, text mining, and sentiment analysis that outline the challenges and techniques associated with each method and emphasizes the benefits of hybrid models for improved accuracy [41]. The review highlights the increasing trend of integrating diverse sources, like social media and news, into prediction models, but also notes limitations such as managing large data volumes and accuracy in sentiment classification [41].

The other researcher also suggested feature-based forecast model averaging (FFORMA), an automated method for weighted forecast combinations based on time series features. However, it also requires consideration of transaction costs and risk strategies, calling for further research to improve its adaptability and responsiveness to market changes [42]-[46]. Along the same lines, the current study also used multiple social media sources to create a hybrid model for predicting stock prices.

Similarly, sentiment analysis research focused on limited parameters and did not use hybrid or ensemble techniques [47]-[49]. The review suggests that better feature selection techniques could improve the accuracy of stock price predictions [48]-[51].

Furthermore, very few studies applied multiple methods to aggregate results [52], [53]. By combining these factors, the proposed decision support system aims to improve the accuracy and precision of stock market predictions for short-term durations. Figure 1 shows the proposed flow chart based on the research gap identified in the literature to address the purpose of the study.

3. METHOD

This study proposes a model to predict short-term stock prices using investor sentiment and technical indicators. It aims to analyze the impact of investor sentiment and micro-blogging on stock prices and identify useful technical indicators for trend identification. The main research objectives are to review sentiment data collection sources, compare text-mining and data-mining techniques, propose a new method to measure the impact of news and investor sentiment, analyze technical indicators, compare different ML techniques, and evaluate the proposed model’s precision, recall, specificity, subjectivity, and accuracy.

3.1. Data and sector selected

The proposed research will be conducted in the banking, pharma, and real estate sectors. These sectors are core, long-lasting, and never-ending in India, representing almost 80% of the economy. They are also the fastest-growing sectors. Bank, Pharma, and Real Estate sector stocks account for a significant portion of mutual fund holdings in most mutual fund schemes, covering approximately 10-15% of holdings depending on the mutual fund’s objective and scheme.

Additionally, these sectors are listed in the top ten significant allocations in the Union Budget allocation of the last three years. For example, in the Union Budget 2021, the Ministry of Health and Family Welfare has allocated 10.35 billion US dollars, while the Ministry of Housing and Urban Affairs has allocated 7.64 billion US dollars. The banking sector represents the country’s overall economy and is a core sector that involves most of the country’s sector development. Therefore, the researcher has selected the banking, Pharma, and real estate sectors for the proposed research work experiments.

Three stocks are selected for the final experiment process in each sector. The stock selection process in each industry is decided by using these three factors: the most extensive capitalization stock in the industry, central holding in the Sensex and Nifty indexes and High liquidity. These parameters are used to select the three different stocks in each sector for the experimental process of the proposed model. Moreover, the stock selection process also considers small-cap, mid-cap, and large-cap factors.

Therefore, Torrent Pharmaceuticals (TORRENTPHARM), Sun Pharmaceutical Industries Ltd (SUNPHARMA), and Biocon (BIOCON) stocks are selected for the experiment in the pharma sector. In the bank sector, the State Bank of India (SBI), HDFC Bank (HDFCBANK), and AXIS Bank (AXISBANK) stocks are selected for the experiment. Godrej Properties Limited (GODREJPROP), DLF Ltd (DLF), and Housing and Urban Development Corporation (HUDCO) stock are selected for the proposed model research experiment process in the real estate sector.

3.2. Proposed model

Figure 1 shows the proposed flow chart based on the research gap identified in the literature to address the study's purpose. It can be described in the following phases:

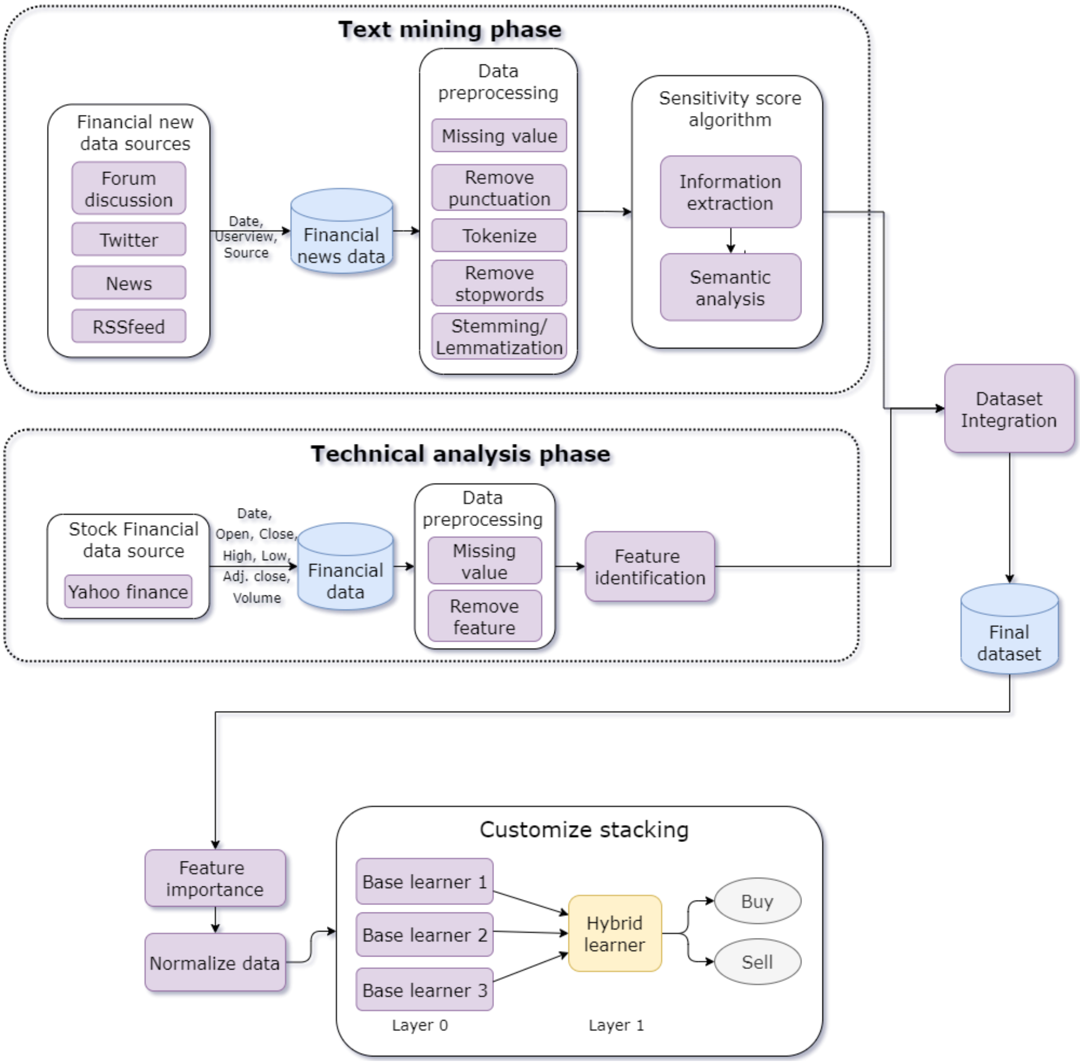

Figure 1. Flow-chart for proposed modelling of equity price prediction
(Source: proposed and prepared by authors)

3.2.1. Text mining phase

Text mining is a subfield of knowledge discovery used to extract hidden patterns and rich information from unstructured and semi-structured textual data. It optimizes daily operations and makes better short-term decisions in various domains. Sentiment analysis is a popular text-mining method that identifies the writer's sentiment and the polarity of views. Figure 2 depicts the text mining process involving three stages: information retrieval, information extraction, and analysis of extracted information using data mining techniques. The extraction stage involves data preprocessing, feature generation, and feature extraction processes. The researcher used the sensitivity score algorithm to find the polarity and semantic class of the user's view.

Text analytics involves retrieving social sentiment documents from various sources, as mentioned in Figure 2, using Python scripts or web crawling tools. The data is preprocessed to remove noise, tokens are generated by creating feature metrics for ML algorithms, and prominent phrases are identified through feature extraction. Financial news datasets are also preprocessed to remove irrelevant data and converted into a structured format using feature generation techniques, such as bag-of-words and TF-IDF.

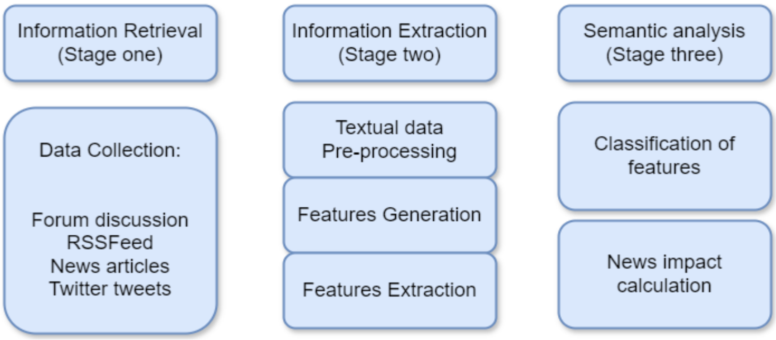

Figure 2. Text mining stages

Sensitivity score is used to calculate token impact and identify dominant tokens and their category value to determine impact. The score is calculated based on several token matches found in three different dictionaries like positive, negative and neutral, by the researcher. All occurrence tokens calculations are done, and then the sensitive score algorithm calculates the score for each day. Further, based on that, each available data score will be calculated and fed into the next stage of the proposed model to understand the influence of sentiment data on the stock for the short-term duration. Moreover, the pseudo-code details for the sentiment score calculation process are given below.

The Pseudo code for the sensitivity score process:

i). Start. ii). Tokenize the user's view (news, document) into a Word vector. iii). Prepare the positive, negative and neutral dictionary containing words(tokens) with their category rank. iv). Calculate the weighted average of all the tokens in each dictionary. v). Check against each word of the user's view to see whether it matches one of the words in the dictionaries. vi). Count the number of matched words occurrence with all the tokens of the dictionaries. vii). Calculate the sensitivity score of each user's view. viii). Classify the news based on sensitivity score. ix). Calculate the impact of news. x). Stop.

3.2.2. Technical analysis phase

Technical analysis is a research technique that uses historical stock price data to identify trading opportunities. This involves understanding the trends and patterns of the stock market, which research analysts can locate. However, computational intelligence computing approaches have emerged as a more efficient and unbiased alternative to technical analysis. Technical analysis is divided into two forms: chart patterns and technical indicators. Chart patterns are shapes that exist in the price chart, and technical indicators analyze the supply and demand of securities.

3.2.3. Data integration and normalizing data

Data collection is the first step in any research work. Open, high, low, close (OHLC) data is used as the base parameter for technical study. Technical data may contain noise and outliers, which should be removed using pre-processing techniques. Data pre-processing helps convert raw data into a structured format that can be used as input for the ML model. Data pre-processing techniques include data cleaning, integration, transformation, and reduction. In the proposed research study, two feature sets are used, which should be combined using data integration techniques. The hypothesis of the Study: Ho: Investors' sentiments and technical indicators-based prediction models do not significantly improve prediction accuracy for the short-term duration.

Stock market forecasting has been a challenging task due to the involvement of technical analysts, which may cause human bias and subjective interpretation. Researchers have developed models based on computational techniques to address this issue. ML algorithms have become famous for finding hidden trends and patterns in the stock market. Various supervised machine-learning approaches are used in forecasting future stock prices, and six supervised machine-learning methods have been chosen for a comparative analysis in the presented research. Statistical techniques and computational techniques are used in stock market predictions. ML techniques significantly improve the accuracy of stock market predictions.

Pseudo code for customized stacking-based prediction model:

i). Start ii). Split the final dataset into training and test sets. The training set is further divided into the K-folds iii). The base learner is fitted on the k-1 parts, and prediction is made for the kth part. Iv). This process continues, and each fold is predicted. v). The base learner model is fitted onto the whole training dataset. vi). The first five steps are to be repeated for the other base learners. vii). The second level of the customized stacking method uses the input parameters of the learners' predictions based on the training set.

viii). Then, the meta-learner takes the results of base learners as input in the training process. ix). Finally, the meta-learner predicts the new data points available in the testing set to check the performance of a customized stacking-based prediction model. x). Stop.

4. RESULTS AND DISCUSSION

Table 1 shows the experimental results using the KNN, DT, ANN, SVM, LR, and RF ML techniques to predict the pricing movement of stocks in different sectors. These six methods are used in Python to create appropriate coding to conclude the signal of trade for particular shares traded on the national stock exchange (NSE). Results are tabulated in accuracy, precision, recall rate and F-score measures for trading signals of buy or sell for a particular company. Accuracy is a ratio of correct predictions to total predictions done by the model. Hence, it is indicative of the success rate of trading. All nine stocks under analysis show top-three accuracy with SVM, LR, and RF prediction methodology.

As shown in Figure 1, text-based sentiment analysis and technical analysis were used for the final data set creation in the proposed Stock market staking-text mining (SMS-TM) model of share trading signal prediction. These are further divided into layer 0 and layer 1, where layer 0 uses SVM, LR, and RF for prediction as base learners one, two and three. Further, layer 1 used the hybrid model to give the final signal of buying or selling stocks that help the traders trade in the equity market to make higher profits than usual.

Table 1. Result comparison of traditional ML models using scikit-learn library

Script name	Methodology	KNN	DT	ANN	SVM	LR	RF
TORRENTPHARM	Accuracy	74.24	83.31	78.79	**87.88**	**86.36**	**83.33**
	Precision	63.33	71.88	65.71	**79.31**	**76.67**	**73.33**
	Recall	76.00	**92.00**	**92.00**	**92.00**	**92.00**	88.00
	F-measure	69.09	80.70	76.67	**85.19**	**83.64**	**80.00**
SUNPHARMA	Accuracy	74.24	77.27	**81.82**	**81.82**	**83.33**	**84.85**
	Precision	**89.29**	83.33	**90.91**	88.57	**93.75**	89.19
	Recall	64.10	76.92	76.92	**79.49**	**76.92**	**84.62**
	F-measure	74.63	80.00	83.33	**83.78**	**84.51**	**86.84**
BIOCON	Accuracy	74.24	75.76	83.33	**86.36**	**86.36**	**86.36**
	Precision	85.29	87.88	**96.88**	**100**	94.44	**97.06**
	Recall	70.73	70.73	75.61	**78.05**	**82.93**	**80.49**
	F-measure	77.33	78.38	84.93	**87.67**	**88.31**	**88.00**
SBI	Accuracy	63.33	80.00	78.33	**88.33**	**90.00**	**83.33**
	Precision	58.33	81.82	76.00	**91.30**	**91.67**	**86.36**
	Recall	53.85	69.23	73.08	**80.77**	**84.62**	**73.08**
	F-measure	56.00	75.00	74.51	**85.71**	**88.00**	**79.17**
HDFCBANK	Accuracy	55.93	67.80	67.80	**76.27**	**79.66**	**71.19**
	Precision	63.64	76.67	**78.57**	**80.00**	**87.10**	76.47
	Recall	60.00	65.71	62.86	**80.00**	**77.14**	**74.29**
	F-measure	61.76	70.77	69.84	**80.00**	**81.82**	**75.36**
AXISBANK	Accuracy	55.00	68.33	66.67	**80.00**	**81.67**	**76.67**
	Precision	45.16	60.00	54.76	**70.00**	**72.41**	**67.86**
	Recall	58.33	62.50	95.83	**87.50**	**87.50**	**79.17**
	F-measure	50.91	61.22	69.70	**77.78**	**79.25**	**73.08**
GODREJPROP	Accuracy	78.33	83.33	73.33	**90.00**	**90.00**	**90.00**
	Precision	80.65	86.67	83.33	**88.24**	**88.24**	**88.24**
	Recall	78.12	81.25	62.50	**93.75**	**93.75**	**93.75**
	F-measure	79.37	83.87	71.43	**90.91**	**90.91**	**90.91**
DLF	Accuracy	67.21	77.05	67.21	**83.61**	**83.61**	**78.69**
	Precision	59.09	66.67	57.69	**75.00**	**75.00**	**67.74**
	Recall	54.17	83.33	62.50	**87.50**	**87.50**	**87.50**
	F-measure	56.52	74.07	60.00	**80.77**	**80.77**	**76.36**
HUDCO	Accuracy	78.33	85.00	80.00	**86.67**	**86.67**	**88.33**
	Precision	78.38	87.88	78.95	**88.24**	**88.24**	**90.91**
	Recall	85.29	85.29	88.24	**88.24**	**88.24**	**88.24**
	F-measure	81.69	86.57	83.33	**88.24**	**88.24**	**89.55**

Source: the author prepared a compilation of accuracy, precision, recall, and F-measures.
In bold, the top three prediction methods are based on particular criteria.

Table 2 shows that the SMS-TM model significantly improves accuracy, precision, recall rate, and F-measures. Hence, the study rejected the null hypothesis. This helps traders achieve better returns in trading. Therefore, using the newly proposed hybrid model, one can take advantage of market inefficiency to create abnormal returns in a short duration. The following section provides a detailed discussion with specific implications and future scope.

Table 2. SMS-TM proposed model results

Script name	Accuracy	Precision	Recall	F-measure
TORRENTPHARM	90.91	85.19	92.00	88.46
SUNPHARMA	90.91	94.74	92.31	93.51
BIOCON	89.39	94.59	85.37	89.74
SBI	90.00	91.67	84.62	88.00
HDFCBANK	86.44	94.12	91.43	92.76
AXISBANK	85.00	91.30	87.50	89.36
GODREJPROP	85.00	93.10	84.38	88.53
DLF	90.16	90.91	83.33	86.96
HUDCO	88.33	93.75	88.24	90.91

Source: the author prepared

5. DISCUSSION, IMPLICATION, AND FURTHER SCOPE

The efficient market hypothesis (EMH) is the foundation for understanding the state of the market, which possibly helps traders create higher returns than average through specific tools and techniques. As per EMH technical analysis, it doesn’t work when the market is weak, fundamental analysis does not work when the market is semi-strong, and even insider information doesn’t work in solid forms of markets. However, to date, creating or achieving strong form markets has not become possible even in developed countries. In developing countries like India, there are vast possibilities of inefficiency in trading platforms response that may create opportunities for traders to earn higher returns than average.

However, creating higher returns requires understanding lots of technical and sentimental data. Today, blogs and RSS feeds have a lot of potential to identify insider information that is the reason for abnormal returns in semi-strong markets. Again, segregating all data and converting data into information, and finally, in the form of a decision to buy-sell, is challenging manually. Hence, the proposed SMS-TM model is the coding of the whole process that directly covers technical and sentiment data to convert it into an accurate prediction of buy-sell calls for traders to create higher returns.

The current paper has used the most efficient six methods for creating prediction models, namely, KNN, DT, ANN, SVM, LR, and RF ML techniques to predict the pricing movement of stocks in different sectors where it was found that SVM, LR, and RF were giving highest accuracy in all the prediction modeling. Hence, that is used as the SMS-TM base model, which uses sample data in two parts. Data for model creation and a second data set to check its predictability and accuracy. It gave better results, and as it is a machine-based model, prediction becomes more accessible for the stakeholders who want to create abnormal returns through the stock market.

The SMS-TM model helps traders create higher returns and makes the market more efficient and stable to prevent the economy’s wealth. In the current study, only nine equity shares from the Indian market are considered a sample, but one can use the same approach for model creation for more equity shares and markets for better results. This is not only for traders, but a current piece of paper also helps the government create a more stable market by motivating various government scheme investments into stock markets.

6. CONCLUSION

The study’s results align with previous work’s expected market prediction to create abnormal returns. SMS-TM is a humble attempt to create a better accuracy model that covers sentiment and technical parameters. As discussed in the results, a hybrid model gives better results than individual techniques-based buy-sell prediction through KNN, DT, ANN, SVM, LR, and RF.

The proposed model is further divided into layer 0 and layer 1, where layer 0 uses SVM, LR, and RF for prediction as base learners one, two and three that’s staking of the model. Further, layer 1 used the hybrid model (that combines layer 0 prediction with text mining from social media that gives another support to trade signal for the particular stock) to provide the final signal of buying or selling stocks that help the traders trade in the equity market to make higher profits than usual.

Hence, one can use this SMS-TM model for trading in the Indian equity market to decide on buying or selling at the appropriate time and can make an average higher return with more than 90% accuracy. This model provides abnormal returns to investors and traders in the equity market. Timing for a prediction model is critical in how much one can create abnormal returns. However, each time a trader decides to trade needs to run this whole method of making a temporary model for identifying the buy or sell signal for the stock of concern.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Dr. Abhishek parikh	✓	✓		✓	✓	✓				✓		✓	✓	
Dr. Chetan Gondaliya	✓	✓	✓			✓	✓	✓	✓					

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

- The data supporting this study’s findings are available on request from the corresponding author through email at f13abhishekp@iima.ac.in.

REFERENCES

[1] S. S. Maini and K. Govinda, “Stock market prediction using data mining techniques,” in *2017 International Conference on Intelligent Sustainable Systems (ICISS)*, Dec. 2017, pp. 654–661, doi: 10.1109/ISS1.2017.8389253.

[2] B. M. Henrique, V. A. Sobreiro, and H. Kimura, “Stock price prediction using support vector regression on daily and up-to-the-minute prices,” *The Journal of Finance and Data Science*, vol. 4, no. 3, pp. 183–201, 2018, doi: 10.1016/j.jfds.2018.04.001.

[3] M. A. Paredes-Valverde, R. Colomo-Palacios, M. D. P. Salas-Zárate, and R. Valencia-García, “Sentiment analysis in Spanish for improvement of products and services: A deep learning approach,” *Scientific Programming*, vol. 2017, pp. 1–11, 2017, doi: 10.1155/2017/2535823.

[4] S. Suman, S. Z. Khan, S. K. Das, and S. K. Chand, “Slope stability analysis using artificial intelligence techniques,” *Natural Hazards*, vol. 84, pp. 727–748, 2016, doi: 10.1007/s11069-016-2453-0.

[5] H. Patel and A. Parikh, “Predicting possible fraud in India using machine learning: An empirical comparison between model for better prediction,” *International Journal of Advanced Science and Technology*, vol. 29, no. 5, pp. 5204–5217, 2020.

[6] A. Parikh, D. Kumari, M. Johann, and D. Mladenović, “The impact of environmental, social and governance score on shareholder wealth: A new dimension in investment philosophy,” *Cleaner and Responsible Consumption*, vol. 8, p. 100101, 2023, doi: 10.1016/j.clrc.2022.100101.

[7] P. Abhishek, “Impact of demonetization on shareholders’ wealth: Case of India,” *Asian Journal of Empirical Research*, vol. 9, no. 9, pp. 217–229, Sep. 2019, doi: 10.18488/journal.1007/2019.9.9/1007.9.217.229.

[8] K. P. S. Raghu and A. Parikh, “Impact of demonetization on B2B and B2C companies stock price and liquidity—evidence from India,” *Global Journal of Accounting and Economy Research*, vol. 3, no. 2, pp. 187–204, 2022.

[9] A. Parikh and M. Baruah, “Exploring overreaction hypothesis for large-cap stocks in the Indian stock market: an empirical evidence of superior returns in Nifty 50,” *Indian Journal of Finance*, vol. 7, pp. 32–40, 2013.

[10] C. Gondaliya, A. Patel, and T. Shah, “Stock prediction using machine learning algorithms with special reference to technical indicators,” in *Smart Innovation, Systems and Technologies*, vol. 248, 2022, pp. 319–327.

[11] D. Kumari and A. Parikh, “A study on wealth management during crisis : an empirical study using downside risk approach in India,” *Indian Journal of Research in Capital Markets*, vol. 10, no. 1, p. 44, Mar. 2023, doi: 10.17010/ijrcm/2023/v10i1/172803.

[12] C. Gondaliya, H. Patel, and A. Patel, “Price prediction with the sectorial effect on the stock market due to COVID-19,” in *Information and Communication Technology for Competitive Strategies (ICTCS 2020): ICT Applications and Social Interfaces*, Singapore: Springer, 2022, pp. 263–270.

[13] C. Gondaliya, A. Patel, and T. Shah, “Sentiment analysis and prediction of the Indian stock market amid the COVID-19 pandemic,” *IOP Conference Series: Materials Science and Engineering*, vol. 1020, no. 1, p. 012023, 2021, doi: 10.1088/1757-899X/1020/1/012023.

[14] M. A. Islam, M. R. Sikder, S. M. Ishtiaq, and A. Sattar, “Stock market prediction of Bangladesh using multivariate long short-term memory with sentiment identification,” *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 13, no. 5, pp. 5696–5706, 2023, doi: 10.11591/ijece.v13i5.pp5696-5706.

[15] A. Ismailova *et al.*, “Forecasting stock market prices using deep learning methods,” *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 14, no. 5, p. 5601, Oct. 2024, doi: 10.11591/ijece.v14i5.pp5601-5611.

[16] T. Wu, D. S. Weld, and J. Heer, “Local decision pitfalls in interactive machine learning: an investigation into feature selection in sentiment analysis,” *ACM Transactions on Computer-Human Interaction*, vol. 26, no. 4, pp. 1–27, 2019, doi: 10.1145/3318142.

[17] W. Huang, Y. Nakamori, and S. Y. Wang, “Forecasting stock market movement direction with a support vector machine,” *Computers and Operations Research*, vol. 32, no. 10, pp. 2513–2522, 2005, doi: 10.1016/j.cor.2004.03.016.

[18] S. Usmani and J. A. Shamsi, “News sensitive stock market prediction: Literature review and suggestions,” *PeerJ Computer Science*, vol. 7, p. e490, 2021, oi: 10.7717/peerj-cs.490.

[19] S. E. A. Ali, F.-W. Lai, P. D. D. Dominic, N. J. Brown, P. B. B. Lowry, and R. F. Ali, “Stock market reactions to favorable and unfavorable information security events: a systematic literature review,” *Computers & Security*, vol. 110, p. 102451, Nov. 2021, doi: 10.1016/j.cose.2021.102451.

[20] V. Shah, "Advancements in deep learning for natural language processing in software applications," *International Journal of Computer Science and Technology*, vol. 4, no. 3, pp. 45–56, 2020.
[21] Q. Lin, "Technical analysis and stock return predictability: an aligned approach," *Journal of Financial Markets*, vol. 38, pp. 103–123, 2018. doi: 10.1016/j.finmar.2017.12.001.
[22] S. I. Song, J. T. Janang, E. Yazi, and F. Morni, "The effects of market strength, information asymmetry, and industrial characteristics on Malaysian firms' CAR during COVID-19 pandemic," *Capital Markets Review*, vol. 30, no. 1, pp. 1–15, 2022.
[23] M. Hanif and A. Sabah, "Stock markets' integration in post financial crisis era: Evidence from literature," *Capital Market Review*, vol. 28, no. 2, 2020.
[24] R. Rahul, S. Sarangi, P. Kedia, and Monika, "Analysis of various approaches for stock market prediction," *Journal of Statistics and Management Systems*, vol. 23, no. 2, pp. 285–293, 2020, doi: 10.1080/09720510.2020.1735262.
[25] C. Y. Liu, S. N. Yao, and Y. J. Chen, "Technical analysis toolkit for neural networks in finance and investing," in *Intelligent Computing: Proceedings of the 2019 Computing Conference, Volume 2*, Cham: Springer International Publishing, 2019, pp. 1170–1174.
[26] H. Patel, S. Parikh, A. Patel, and A. Parikh, "An application of ensemble random forest classifier for detecting financial statement manipulation of indian listed companies," in *Advances in Intelligent Systems and Computing*, vol. 740, 2019, pp. 349–360.
[27] A. Bhardwaj, Y. Narayan, dan M. Dutta, "Sentiment analysis for Indian stock market prediction using Sensex and Nifty," *Procedia Computer Science*, vol. 70, pp. 85–91, 2015, doi: 10.1016/j.procs.2015.10.043.
[28] M. Vicari dan M. Gaspari, "Analysis of news sentiments using natural language processing and deep learning," *AI & Society*, vol. 36, no. 3, pp. 931–937, 2021, doi: 10.1007/s00146-020-01111-x.
[29] N. Chatelais, A. Stalla-Bourdillon, dan M. D. Chinn, "Forecasting real activity using cross-sectoral stock market information," *Journal of International Money and Finance*, vol. 131, p. 102800, 2023, doi: 10.1016/j.jimonfin.2023.102800.
[30] T. Shah and A. Parikh, "Does the number of holdings in a risk parity portfolio matter?," *Journal of Asset Management*, vol. 20, no. 2, pp. 124–133, Mar. 2019, doi: 10.1057/s41260-019-00110-y.
[31] M. Sykora, "Engineering social media driven intelligent systems through crowdsourcing: Insights from a financial news summarisation system," *Journal of Systems and Information Technology*, vol. 18, no. 3, pp. 255–276, 2016, doi: 10.1108/JSIT-03-2016-0019.
[32] C. Y. Lin dan J. A. L. Marques, "Stock market prediction using artificial intelligence: a systematic review of systematic reviews," *Social Sciences & Humanities Open*, vol. 9, p. 100864, 2024, doi: 10.1016/j.ssaho.2024.100864.
[33] P. Chhajer, M. Shah, dan A. Kshirsagar, "The applications of artificial neural networks, support vector machines, and long–short-term memory for stock market prediction," *Decision Analytics Journal*, vol. 2, p. 100015, 2022, doi: 10.1016/j.dajour.2021.100015.
[34] M. M. Alshater, I. Kampouris, H. Marashdeh, O. F. Atayah, dan H. Banna, "Early warning system to predict energy prices: the role of artificial intelligence and machine learning," *Annals of Operations Research*, vol. 345, no. 2, pp. 1297–1333, 2025, doi: 10.1007/s10479-022-04908-9.
[35] H. Hewamalage, C. Bergmeir, dan K. Bandara, "Recurrent neural networks for time series forecasting: current status and future directions," *International Journal of Forecasting*, vol. 37, no. 1, pp. 388–427, 2021, doi: 10.1016/j.ijforecast.2020.06.008.
[36] T. Januschowski *et al.*, "Criteria for classifying forecasting methods," *International Journal of Forecasting*, vol. 36, no. 1, pp. 167–177, 2020, doi: 10.1016/j.ijforecast.2019.06.001.
[37] D. Kumar, P. K. Sarangi, dan R. Verma, "A systematic review of stock market prediction using machine learning and statistical techniques," *Materials Today: Proceedings*, vol. 49, pp. 3187–3191, 2022, doi: 10.1016/j.matpr.2020.11.399.
[38] A. W. Li dan G. S. Bastos, "Stock market forecasting using deep learning and technical analysis: a systematic review," *IEEE Access*, vol. 8, pp. 185232–185242, 2020, doi: 10.1109/ACCESS.2020.3030226.
[39] M. J. Page *et al.*, "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews," *BMJ*, vol. 372, p. n71, 2021, doi: 10.1136/bmj.n71.
[40] N. Pinto, L. da Silva Figueiredo, dan A. C. Garcia, "Automatic prediction of stock market behavior based on time series, text mining and sentiment analysis: a systematic review," dalam *Proceedings of the 2021 IEEE 24th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, 2021, pp. 1203–1208, doi: 10.1109/CSCWD49262.2021.9437732.
[41] A. Samitas, E. Kampouris, dan D. Kenourgios, "Machine learning as an early warning system to predict financial crisis," *International Review of Financial Analysis*, vol. 71, p. 101507, 2020, doi: 10.1016/j.irfa.2020.101507.
[42] P. Montero-Manso, G. Athanasopoulos, R. J. Hyndman, dan T. S. Talagala, "FFORMA: feature-based forecast model averaging," *International Journal of Forecasting*, vol. 36, no. 1, pp. 86–92, 2020, doi: 10.1016/j.ijforecast.2019.02.011.
[43] S. Smyl, "A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting," *International Journal of Forecasting*, vol. 36, no. 1, pp. 75–85, 2020, doi: 10.1016/j.ijforecast.2019.05.012.
[44] A. Safari dan M. A. Badamchizadeh, "DeepInvesting: Stock market predictions with a sequence-oriented BiLSTM stacked model–A dataset case study of AMZN," *Intelligent Systems with Applications*, vol. 24, 2024, doi: 10.1016/j.iswa.2022.200439.
[45] F. Ricchiuti dan G. Sperlí, "An advisor neural network framework using LSTM-based informative stock analysis," *Expert Systems with Applications*, vol. 259, p. 125299, 2025, doi: 10.1016/j.eswa.2023.125299.
[46] S. A. S. Syed, "Stock market in the age of COVID-19: Mere acclimatization or Stockholm syndrome?," *The Journal of Economic Asymmetries*, vol. 25, p. e00245, 2022, doi: 10.1016/j.jeca.2021.e00245.
[47] Y. Huang, C. Deng, X. Zhang, dan Y. Bao, "Forecasting of stock price index using support vector regression with multivariate empirical mode decomposition," *Journal of Systems and Information Technology*, vol. 24, no. 2, pp. 75–95, 2022, doi: 10.1108/JSIT-06-2021-0122.
[48] X. Han dan D. Yao, "Exploration of portfolio selection and risk prediction in financial markets based on the SVM algorithm," *International Journal of Information Technology and Web Engineering*, vol. 18, no. 1, pp. 1–16, 2023, doi: 10.4018/IJITWE.2023010101.
[49] A. Huang, W. Wu, dan T. Yu, "Textual analysis for China's financial markets: a review and discussion," *China Finance Review International*, vol. 10, no. 1, pp. 1–15, 2020, doi: 10.1108/CFRI-11-2018-0120.
[50] Z. Amry and B. H. Siregar, "ARIMA model selection for composite stock price index in Indonesia stock exchange," *International Journal of Accounting and Finance Studies*, vol. 2, no. 1, p. p31, May 2019, doi: 10.22158/ijafs.v2n1p31.
[51] G. Sornavalli, G. Angelin, dan N. H. Khanna, "Intelligent forecast of stock markets to handle COVID-19 economic crisis by modified generative adversarial networks," *The Computer Journal*, vol. 65, no. 12, pp. 3250–3264, 2022, doi: 10.1093/comjnl/bxac043.
[52] X. Wang, X. Pan, T. Yang, J. Xie, dan M. Tang, "Aspect-based sentiment analysis using interaction matrix and global attention neural network," *The Computer Journal*, vol. 66, no. 5, pp. 1167–1183, 2023, doi: 10.1093/comjnl/bxac128.
[53] W. He, X. Tian, Y. Chen, dan D. Chong, "Actionable social media competitive analytics for understanding customer experiences," *Journal of Computer Information Systems*, vol. 56, no. 2, pp. 145–155, 2016, doi: 10.1080/08874417.2016.1164004.

BIOGRAPHIES OF AUTHORS

Dr. Chetan Gondaliya is an associate professor and head of the Centre of Excellence at the Swarrnim School of Computing and IT, Swarrnim Startup and Innovation University in Gujarat, India. His research interests include stock market prediction, sentiment analysis, artificial intelligence, natural language processing, agriculture, and information technology. He has more than 12 national and international publications in his credit. His research interests include stock market prediction, sentiment analysis, artificial intelligence, natural language processing, agriculture, and information technology. He can be contacted at email: chetan.p.gondaliya@gmail.com.

Dr. Abhishek Parikh has over 17 years of experience in teaching, training, research, and consultancy. He is a Professor, Principal and Associate Dean at the School of Liberal Arts and Management, P P Savani University. He has more than 35 national and international publications to his credit. He has also taken different workshops, seminars, conferences, MDPs, and FDPs as a resource person at various institutes of national and international repute, such as IIMK, IIMJ, NU, GTU, JIM, SGGU, GGE, and more. He can be contacted at email: f13abhishekp@iima.ac.in.