Sentiment Analysis Using Machine Learning for Forecasting Indian Stock Trend: A Brief Survey

A. S. Dash, U. Mishra

MIT College of Management, MIT Art, Design & Technology University, Pune, India

ABSTRACT

Due to new technical advances, the machine can think as a person-investor and express its reaction to readily available financial information. Forecasting models for the Indian stock market can be developed based on the analysis of these sentiments. The **purpose** of the study is to identify gaps in existing approaches to the analysis of sentiments and models of forecasting trends in the Indian stock market, which can improve the accuracy of the prediction of the dynamics of Indian stocks. The paper presents an overview of the literature on the analysis of sentiments of financial information using lexical methods, machine learning methods and forecasting for the Indian stock market based on sentiment analysis data. The scientific works, conference reports, dissertations, books and articles published by scientists for the period from 2015 to 2021 are considered. The datasets published in Indian Stock Exchanges suggest increasing participation of retail investors in the Indian Stock market in recent times. To help investors in decision-making, various prediction models are available based on the financial information. The results of the survey showed that investors' attitudes based on the microeconomic and macroeconomic information associated with stocks influence the movement of the stock price. Therefore, forecasting a future trend or price requires a sentiments analysis based on available financial information. It was **concluded** that using machine learning to extract sentiment from financial data allows for more accurate forecasts than sentiment analysis based on vocabulary. The results of this study can be useful for students and new professionals in the field of financial information data analysis and stock market predictions who want to get connected with this area, identify problem concerns, and develop models for predicting decision-making.

Keywords: sentiment analysis; stock market; prediction; machine learning; decision making; trend analysis

INTRODUCTION

Investors in the Indian stock market invest their money into retail investors, foreign institutional investors, mutual funds, insurance funds, pension funds, banks, and so on. The investment decisions on the stock market are to gain good returns with the movement of stock prices and dividends. The movement of the price of a stock depends on several factors, such as company performances, announcements, microeconomic conditions, macroeconomic environment, the sentiment of investors towards the stocks, and any new information associated with the stock.

The investors take the long positions with an upward movement or trend of the stock price, so that they gain from the uptrend and with a downward movement or trend, the investor takes the short positions to cover their loses. The correct investment decisions made by investors on stock markets during an upward trend and the downward trend will increase the potential of earning better and adjust the risks in a better way. Many times, due to inaccurate considerations while investing in the stock market, investment decisions go wrong, bringing enormous losses to the investors. The retail investors bear the brunt due to inadequate information on market trends.

To provide a correct decision and to reduce their risk, the prediction of stock price trends is necessary for the short- or long-term based on the investment horizon.

The investor sentiment towards the stock provides a good input to predict the stock price trend using machine learning algorithms. The sentiment analysis of the financial information available related to companies identifies the behaviour of human sentiment towards the stock information available currently. Sentiment analysis is a popular language processing technique where the polarity of the textual data is determined. The sentiment analysis is performed in 2 ways: the

classical Lexicon-Based Approach and the Machine Learning-Based Approach.

This research aims to find out major studies related to Sentiment Analysis processes for financial market information through a deep literature survey and explore the use of machine learning services for sentiment analysis of financial market information for the Indian market, which helps predict the market trend.

THE OBJECTIVE OF THE STUDY

The objectives of this study of a literature review on sentiment analysis through machine learning for Indian financial market are following:

1) To find out various research works completed in the field of sentiment analysis through machine learning for Indian financial markets for the period 2015 to 2021.

2) To find out the impact of sentiment analysis on the price movement of a stock in the Indian stock market.

3) To find out the prediction models used for predicting the trend of stock price movement in the Indian stock market using sentiment analysis.

METHODOLOGY

This research is descriptive research, where the study of literature is mainly based on the available secondary sources such as previous research papers, conference papers, journal papers, past PhD theses, books, online blogs, and articles by various research scholars and academicians for the period of 2015 to 2021. The research for this work is done in both online and offline modes, where literature was identified using specific criteria and word searches. The keywords used to filter out the right kinds of literature are "Sentiment Analysis through Machine Learning", "Sentiment Analysis on Indian Financial Market Information", "Impact of Sentiment Analysis on Indian Financial Market Information".

The secondary kinds of literature were searched based on the specified keywords. In this research, more than 140 papers were scanned, more than 20 PhD theses were referred, and more than 50 various blogs along with online articles were studied. The studied kinds of literature were categorized based on their relevance to the topic, contribution of knowledge, and contribution of research work related to sentiment analysis on the Indian stock market. The pieces of literature related to sentiment analysis on foreign markets were kept in low- and medium-importance categories to focus only on the Indian market.

INDIAN STOCK MARKET

The Indian stock market is one of the oldest stock markets in Asia. The Bombay Stock Exchange (BSE) was established in 1875 as "The Native Share and Stockbrokers Association".[1] After the liberalization of the Indian market in 1991, the National Stock Exchange (NSE) was incorporated in 1992, and it was recognized as a stock exchange by the "Security Exchange Board of India" in April 1993.[2]

As per Bombay Stock Exchange market capitalization data published in October 2021, the total market capitalization was 259 Lakh Crore.[3]

As per the Market Statistics — October 2021 report by the WFE Statistics team, the market capitalization of NSE is 3.4 trn USD, or 252 Lakh Crore Rupees as of August 2021.[4]

The companies are listed on either one or both stock exchanges. In BSE around 5 213 companies were listed as of August 2021[5] and in NSE around 1 920 companies were listed as of August 2021.[6]

TYPE OF INVESTORS

Various investors invest in these listed companies through the NSE or BSE. Private Indian promoters

---

[1] BSE History & Milestones — Bombay Stock Exchange. 2021. URL: https://www.bseindia.com/static/about/History_Milestones.html (accessed on 26.06.2022).

[2] NSE History & Milestones — National Stock Exchanges. 2021. URL: https://www.nseindia.com/national-stock-exchange/history-milestones (accessed on 26.06.2022).

[3] BSE Market Capitalization Report — Bombay Stock Exchange. 2021. URL: https://www.bseindia.com/markets/equity/EQReports/AllIndiamktcap_Histori.aspx (Accessed on 26.06.2022).

[4] BSE History & Milestones — Bombay Stock Exchange. 2021. URL: https://www.bseindia.com/static/about/History_Milestones.html (accessed on 26.06.2022).

[5] BSE Market Capitalization Report — Bombay Stock Exchange. 2021. URL: https://www.bseindia.com/markets/equity/EQReports/AllIndiamktcap_Histori.aspx (accessed on 26.06.2022).

[6] NSE History & Milestones — National Stock Exchanges.2021. URL: https://www.nseindia.com/national-stock-exchange/history-milestones (accessed on 26.06.2022).

hold around 34.6% of the total market capitalization of NSE-Listed companies, followed by Foreign Institutional Investors, holding 21.7%, foreign promoters holding 9.7%, and retail investors holding 9% as of December 2020 (*Fig. 1*).[7]

The shares are being traded on these exchanges daily. NSE's turnover for March 2021 was 13.9 trn rupees.[8] The motive of the transactions is to gain monetary benefit. In these transactions, Foreign Institutional Investors (FIIs), Domestics Institutional Investors (DIIs) Mutual Funds, Corporates, Proprietary Traders, Individual Investors, and others such as Trusts, Partnership firms, VC Funds, etc. are involved. Traditionally, qualified investors, such as FII, and DII participated much more than individual investors. However, in the last 6 years, this trend has changed.

In the financial year 2021, individual investors, such as individual domestic investors, NRIs, sole proprietorship firms, and HUFs, will account for 45% of the total turnover in the cash segment of the NSE. As per the report, NSE has added 90 Lakh new investors in the current fiscal year. More individual investors are now entering the Indian stock market to participate in stock market trading (*Fig. 2*).

STOCK PRICE MOVEMENT

Stock prices change based on the demand and supply of shares on the exchanges. If the demand for the stock exceeds the supply, the price of the stock rises; similarly, if the supply of the stock exceeds the demand, the price of the stock falls. The movement of the price of a particular stock can behave the same way for a few days, a few weeks, or even months based on demand and supply.

The direction of price movement is termed a trend [1]. The trend can be upward or downward. An upward trend is formed when the price movement makes higher swing highs and higher swing lows. A downward trend is formed when the price movement makes lower swing lows and lower swing highs [2]. The trends are measured

for the short-term, medium-term or intermediate-term, and long term. The short-term trend can be the price movement in the same direction for a few days to a few weeks; the medium-term trend can be the price movement in the same direction for a few weeks to a few months; and the long-term trend can be the price movement in the same direction from a few months to a few years [3].

TYPE OF STOCK PRICE ANALYSIS

According to the Efficient Market Hypothesis, the price of a stock reflects all the information available to the market, and it is impossible to beat the market [4]. However, the EMH is highly controversial, as modern financial theory suggests no market is perfectly efficient, and thus stock prices do not always accurately reflect their true value. The Indian stock market returns are not completely random. It exhibits a weak form of market efficiency, so the prediction of the true value of stock price exists [5]. Investors try to perform various price analyses of a particular stock before deciding to buy or sell that stock. Predominately, two types of stock price analysis — fundamental analysis and technical analysis — are performed by investors. Fundamental analysis is to identify the stock's correct value by examining various economic factors from the micro to macro level.

On the other hand, technical analysis is the study of historical price movements of securities and the patterns through charts and various indicators to identify the correct price of the security and forecast the future movement.

SENTIMENT ANALYSIS

As per the theory of behavioural finance, Noise in the form of information makes the market inefficient. Also, testing theories in the financial market is very difficult due to the presence of Noise. Under the influence of Noise, traders or investors behave irrationally and take the stock price away from the true price level. However, the true price of stock returns in the long run [6].

The stock price tends to be diverse from its true value under the influence of noisy signals present in the financial market. The traders act on these signals and move the true price of a stock in a different direction. So, the investors are subject to sentiments; this sentiment

---

Fig. 1. **Types of Investors Holding in Percentage the Ownership of NSE-Listed Companies as of December 2020**
Source: Market Pulse-A Monthly review of Indian economy and markets. 2021. URL: https://static.nseindia.com//s3fs-public/inline-files/Market_Pulse_April_2021.pdf (accessed on 28.06.2022).
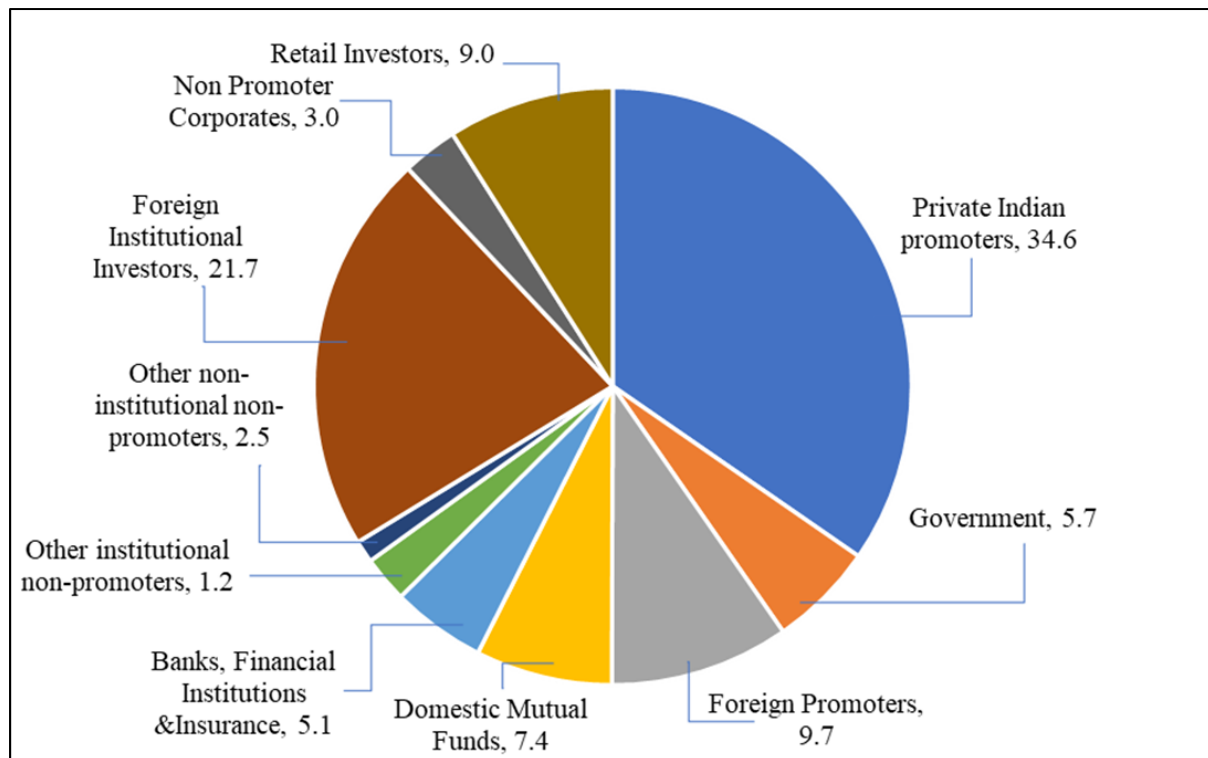

Fig. 2. **Types of Investors Participate in the Capital Market at NSE (in %)**
Source: Market Pulse-A Monthly review of Indian economy and markets. 2021. URL: https://static.nseindia.com//s3fs-public/inline-files/Market_Pulse_April_2021.pdf (accessed on 28.06.2022).
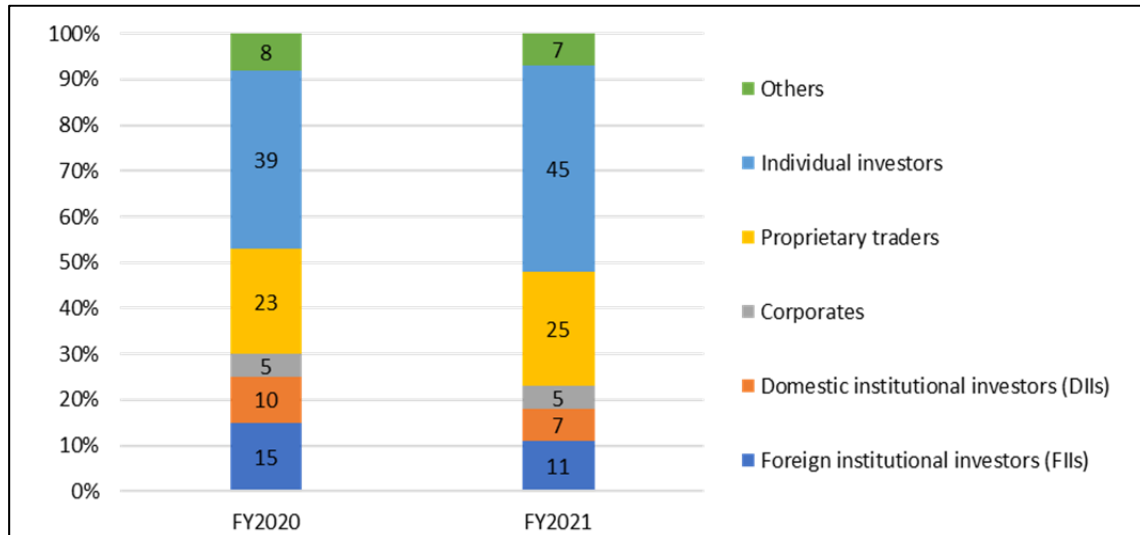
is their belief about the future cash flow of the security, which is not justified by the available information [7].

Based on the research by Baker & Wurgler, it has been clear that investor sentiments affect the stock price movement. In their work, they have proposed a top-down approach to measure investor sentiment [8, 9]. Measuring investor sentiment is not straightforward; however, using imperfect proxies, they created a model

to measure the sentiment. In their model, they have listed out various proxies such as investor psychology to respond the corporate news, trade volumes, mutual fund investments, announcement of dividends, implied volatility of stock options, listing day returns on initial public offerings (IPOs), the volume of initial public offerings, new equity issues, and insider trading information. Based on these proxies, they provided the sentiment index as

$$SENT = -0.23CEFD + 0.23TURN + 0.24NIPO + 0.29RIPO + 0.32PDND + 0.23S.$$

In this equation, they have used 6 proxies or factors to define the sentiment. These 6 factors are the closed-end fund discount (CEFD), detrended log turnover (TURN), number of IPOs (NIPO), first-day return on IPOs (RIPO), dividend premium (PDND), and equity share in new issues (S). The major challenge of this model was characterizing and measuring uninformed investor sentiment and the variation in investor sentiment over time.

On a similar principle, various studies were conducted on the Indian stock market to identify the relationship between investor sentiment and stock market volatility. Investor sentiment plays a vital role in the Indian stock market's volatility. As per the study, investors are more responsive to negative news than positive. The negative sentiment plays a greater role in the volatility of the Indian market [10].

To measure the linkage between investor sentiment and stock market volatility, similar studies have been performed on the Indian stock market along with the world stock market. Jyoti & Jitendra, in their research, went ahead to assess the predictability the asset volatility in the Indian stock market using investor sentiment. To measure investor sentiment through their model, they provided six macro-economic factors such as the Index of industrial production (IIP), short-term interest rates as Treasury bill rates (TBR), term spread (difference between long-term bond yield and Treasury bill rate), the exchange rate (EX), wholesale price index (WPI) and foreign institutional investments (FII) and four market-symmetric factors such as four market-wide systematic factors, market risk premium (Mkt), the premium on the portfolio of small stocks relative to large stocks (SMB), the premium on the portfolio of high book/market stocks relative to low book/market stocks (HML) and momentum factor (WML). Their study concluded that investor's sentiment does predict the volatility of assets in the Indian stock market. In this research, investor sentiment is predominately derived from the available quantitative data [11].

The sentiment is human behaviour, and any information, whether qualitative or quantitative, has some influence on the sentiment value. In the past few decades, as information availability has tremendously increased, many researchers have worked towards an understanding of investors' sentiment using both qualitative and quantitative data. When any new information is available, it has some degree of impact on the sentiment of the investor [12].

In recent times, much research has been conducted to identify any causal relationship between financial information and the impact on Indian stock markets. P. Misra in his research, identified the relationship between BSE Sensex and macroeconomic variables such as the Index of Industrial Production (IIP), inflation, the rate of interest, the price of gold, the rate of exchange, FII, and supply of money. He also confirms that there exists a long-term causality between these macroeconomic variables and BSE Sensex [13].

Some of the researchers conducted the study on individual variables. As Foreign Institutional Investors are one of the largest players in the Indian stock market, their net investment is positively influenced by the NIFTY returns [14].

Financial information related to microeconomic and macroeconomic variables are announced and published through national newspapers, company websites, or NSE and BSE websites. These announcements impact the sentiment of investors towards the stock. There is a significant impact of the announcement of a new product launch or approvals or decisions on the company's share prices [15]. So abnormal returns are generated on the event day. The sentiment from the news event has some time limit on the investor. The effect of positive or negative sentiments lasts for certain days on the stock price movement from the event date [16].

The investor sentiment factor based on financial news adds significantly to the traditional asset pricing model [17]. The news articles were collected from the published data of the Guardian Newspaper and

sentiment analysis was conducted to understand the impact on the London stock market. Based on the analysis, it was found that the sentiment metric influences the volatility of the London Stock Exchange Index [18].

Financial information published on the listed stock exchange, newspaper, or microblogging sites is unstructured. Chan & Chong, in their research, published a model on how to extract insights from unstructured data for sentiment analysis [19]. They also identified that this information contains noise, which needs filtering. A detail filtering technique to reduce the noise present in financial news sentiments was discussed in the research carried out by M.W. Uhl [20]. To understand the sentiment of the investor when new financial information is available, sentiment analysis is needed on the available financial news.

As per the Oxford dictionary definition, sentiment analysis is the process of computationally identifying and categorizing opinions expressed in a piece of text, especially to determine whether the writer's attitude towards a particular topic, product, etc. is positive, negative, or neutral.[9] Sentiment analysis is a popular language processing technique where the polarity of the textual data is determined.

Sentiment analysis is a branch of Natural Language Processing (NLP) for analysing public opinion [21]. The sentiment analysis is performed in 2 ways, as shown by F.Z. Xing et al. [22]. One is the classical Lexicon-Based Approach, and the other is the Machine Learning-Based Approach.

In the Lexicon-Based Approach, dictionaries of words are mapped with emotional polarities, such as positive, negative, and neutral. Then these words are matched to the input data to calculate the overall polarity of the data [23]. *Figure 3* represents the flow chart for a Lexicon-Based Sentiment Analysis Model.

MACHINE LEARNING APPROACH

In the Machine-Learning Approach, various machine learning algorithms are used to identify the polarity of the textual data. Machine-Learning is a subset of Artificial Intelligence, where complex algorithms are programmed to replicate the function of the Human Brain in such a way that they can solve complex problems in the same way the human brain solves.

Machine learning is an application of artificial intelligence that provides the system with the ability to automatically learn from the training set of data and improve on the experience it has gained from the training data without being explicitly programmed for it. Machine learning focuses on the learning capabilities of computer programs that can access data and use it to learn for themselves what the human brain does.

In the machine learning approach, two datasets are used, where the first one is used to train the model and the second dataset is used to Test the model (*Fig. 4*).

In the past, several studies have been conducted to test sentiment analysis from various social media and news sources using machine learning algorithms. It has been concluded that the accuracy and efficiency of sentiment analysis using the machine learning approach are better than the Lexicon-based approach [21].

There are various challenges involved in creating a system to analyse the sentiment, using the machine learning approach. Infrastructure is one of the major challenges, as a huge amount of data is required to train the machine learning model and then test to get the desired accuracy. The system should be able to store and process this data. Another challenge is the use of machine learning algorithms and desired expertise in that field to create these models. As a result, small, medium-sized organizations or individuals do not have the necessary resources to build some of these sentiment analysis systems [26].

To overcome these challenges, we need to identify major cloud-based AI platform services that offer sentiment analysis through Application Programming Interfaces using their powerful supercomputers in a very cost-effective way. Following are a few major technology solutions currently available:

• IBM Watson Natural Language Understanding Services;

• Microsoft Azure Cognitive Services;

• Google Cloud Natural Language AI;

• Amazon Comprehend.

These major technology companies are providing Application Programming Interfaces (APIs), which enable developers or individuals to use these interfaces in their programs and receive the Sentiments analysis

---

[9] Online Oxford Dictionaries. 2019. URL: https://www.oxfordlearnersdictionaries.com/ (accessed on 28.06.2022).
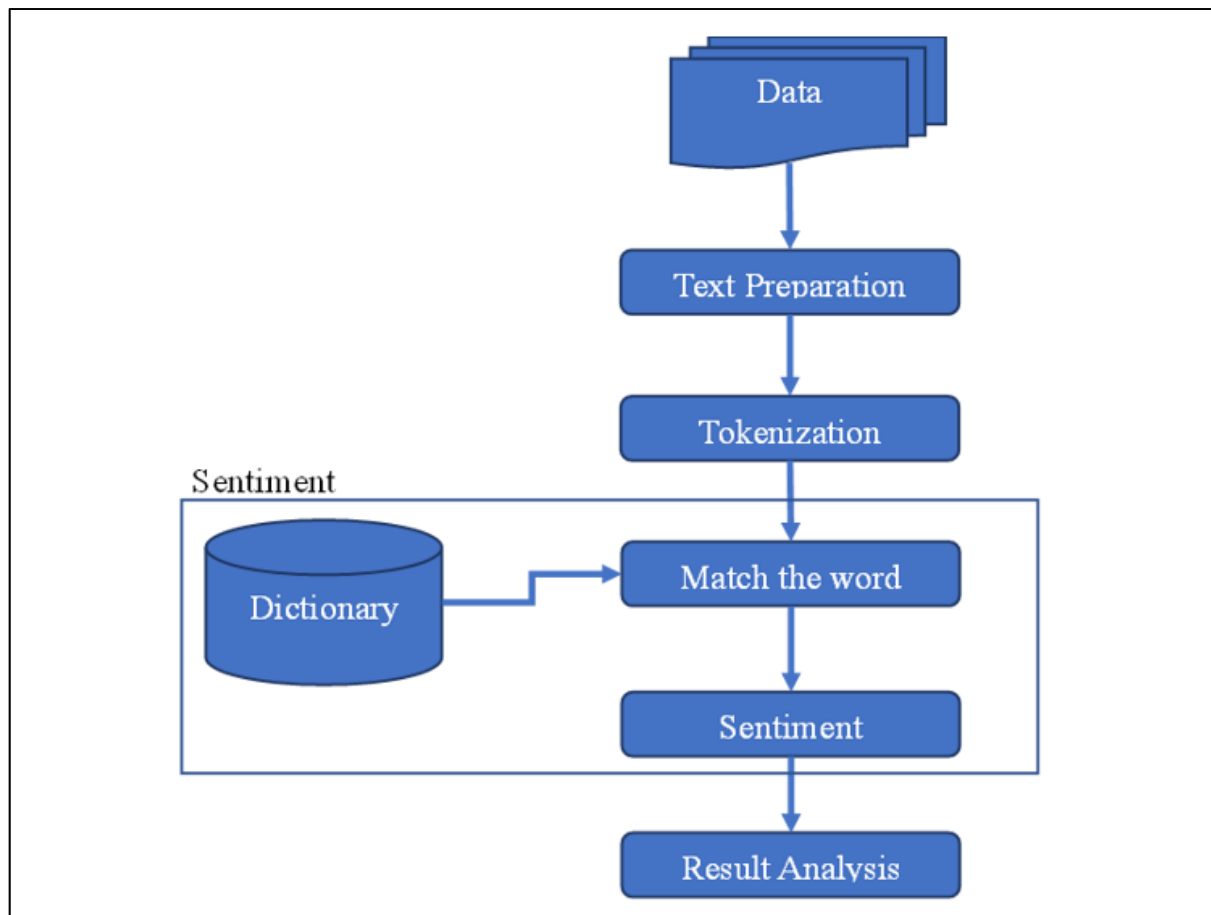
Fig. 3. **Lexicon-Based Sentiment Analysis Process on Financial News**
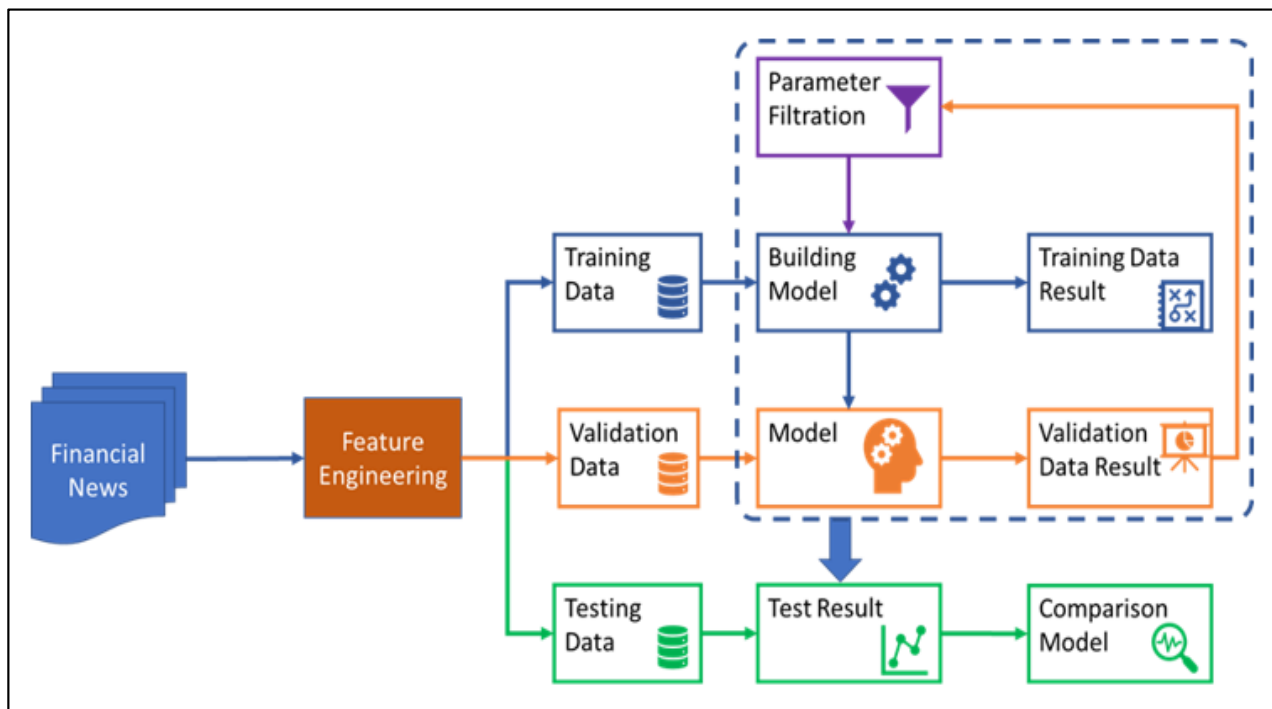*Source:* H.A. Shehu et al. [24].

Fig. 4. **Machine Learning-Based Sentiment Analysis Process**
Source: E. Alpaydin [25].

Summary of the Technologies and the Overall Accuracy Level

Cloud Platform	Technology	Output	Accuracy Level
IBM Cloud1	Natural Language Understanding	Score inside the interval [−1,1]	97%
Microsoft Azure2	Text Analytics	Score inside the interval [0,1]	94%
Google Cloud2	Natural Language	Score inside the interval [−1,1]	90%

Source: Based on the de Las Heras-Pedrosa C. et al. [27], Carvalho A., Xu J. [28].

results for the input data. This eliminates the infrastructure challenges of creating the models and deep knowledge of technical programming.

The IBM Watson Natural Language Understanding has an overall accuracy of 97% for natural language processing. "Its performance has been compared with other systems as well as humans, and in either case, the result is very satisfactory" (see *Table*) [27].

The accuracy testing is conducted by Carvalho & Xu. In their research, they tested the accuracy of each of the above systems with 14 605 tweets and 3,209 Facebook posts. As per the result, they concluded that IBM Watson's accuracy is better than other systems [28].

Figure 5 depicts the process flow for sentiment analysis of texts using the IBM Watson Natural Language understanding API services.

Various documents or web contents related to the financial market or stock-specific news feed into IBM Watson through the Natural Language understanding API services. Then the result has been provided with a sentiment score in a range of −1 to 1, where −1 represents Negative, 0 represents Neutral and 1 represents positive.

The IBM Watson Natural Language platform also provides additional flexibility through custom training of specific domains to the system, where custom training data in an Excel file can be fed to the system to increase the accuracy of the system for a specific domain. With this approach, additional documentation is provided with the sentiments provided to the machine to learn more about the specific domain as humans do.

After the development of Machine Learning Algorithms and related concepts, various researchers and academics turned their focus on predicting the stock price using this technology. Most of the research was conducted to predict stock prices based on their previous price through various machine learning Algorithms such as long- and short-term memory, Artificial Neural network, etc. [29–32].

In the recent study by M. Obthong et al. [33], they did a survey of literature on Machine learning algorithms and related models to predict the price of the stock market based on previous prices. They found the accuracy of these systems ranges from 55% to 65%. They concluded to increase the accuracy of the prediction models, not only the previous price of the stock but also additional information such as Sentiment towards the stock will be needed.

The news information related to stocks provides sentiment towards the stock price, and it helps in the price movement reflected on the stock price. With recent technological developments in natural language understanding, the sentiment analysis of these news stories provides a vital input to predicting the stock price.

The results from the work by H. Rich et al. [34] are quite encouraging, where they created a model to use the New York Times published information related to renewable energy sector companies to extract sentiments, and then they used the stock market data in their machine learning algorithms to predict the renewable energy index price. With their approach, they have achieved an accuracy of 75% in the predictability of the NASDAQ renewable energy index trend. Similar research was performed by G.G.-R. Wu et al. [35] to understand the stock market returns from the Taiwan stock exchange using the new-based sentiment analysis
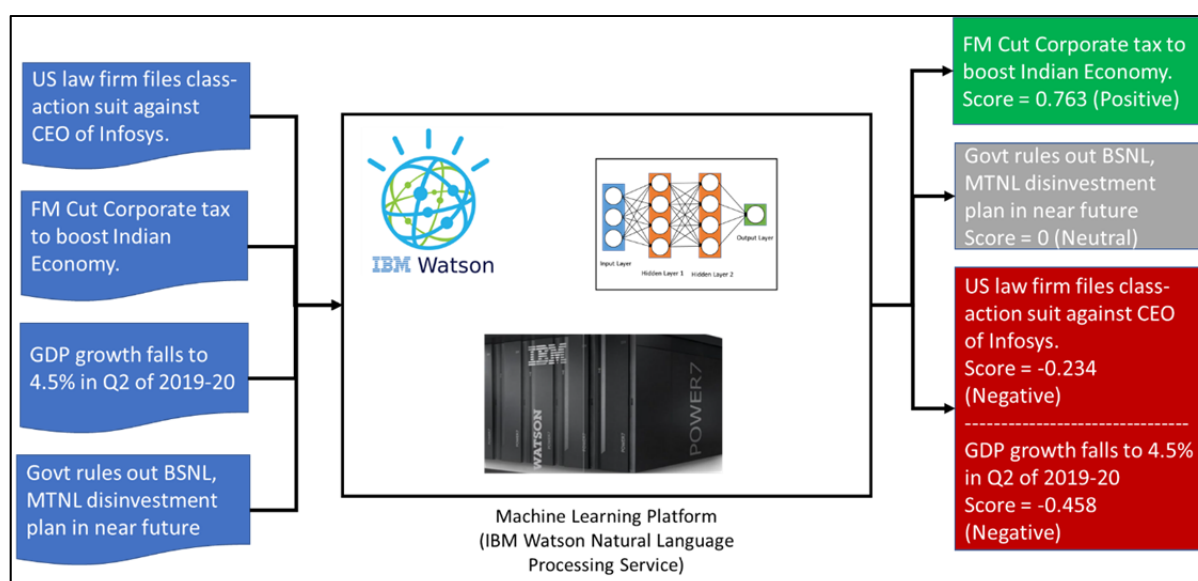
Fig. 5. **Data Flow for Sentiment Analysis of Texts Using the IBM Watson Natural Language Understanding API Services**
Source: A. Carvalho, J. Xu [28].

and market data. They concluded that the news variable provides useful information for predicting the Taiwan market returns.

REVIEW OF SENTIMENT ANALYSIS FOR INDIAN STOCK MARKET

Researchers are continuously working on the technological advancement of sentiment analysis using AI-driven systems and machine learning algorithms to provide better accuracy in predicting the price movement of stocks. Though limited studies were performed on Indian stock markets compared to the US and world stock markets, the Indian stock market behaves very poorly. Market efficiency and prediction opportunities exist in the Indian stock market [5].

It is evident from the sentiment of investors on the news of any new product announcement for the BSE 500 index companies that there is a significant impact of the news of new products on the stock price. With the information of the new product launch, the trend of stock price changes and abnormal returns are generated on the event day [15].

Macroeconomic factors such as Industrial Production (IIP), inflation, the rate of interest, the price of gold, rate of exchange, FII, and supply of money have an impact on the BSE Sensex index price movement. There exists a causal relationship between the information related to macroeconomic variables and the price movement of BSE Sensex. Also, the impact of macroeconomic variables such as foreign portfolio investors' sentiment influences positively the movement of NIFTY returns [13, 15].

R. Yadav et al. [23] in their study, created an event-based sentiment analysis model to predict the Indian stock market prices. They have implemented the Lexicon-based sentiment analysis model to predict the sentiment for the news feed. They found the model provided an additional aid while deciding on an investment.

The research by N. Rani et al. [36] on the NIFTY 50 index stocks of the Indian stock exchange provides a significant relationship between the index price movement and the sentiment score. In their study, they collected new information from Twitter and published news, using an available machine learning model to predict the sentiment. They measure the NIFTY 50 index returns based on the sentiment score. They concluded that there is a significant relationship between the sentiment score and the NIFTY 50 index return on a 10-day moving average.

CONCLUSION

From our research and review of articles, it is evident that the Indian stock market behaves

weakly in terms of market efficiency and prediction opportunities. In the previous research work, various prediction models were created to predict the trend or price of stocks in the global as well as Indian stock markets using machine learning algorithms. The majority of this research was aimed at using the past stock price alone, which is not sufficient to predict the trend or price accurately. They do not provide the complete sentiment of the investors towards that stock price movement.

From the studies, it has been concluded that the sentiments of investors from the microeconomic and macroeconomic information related to the stock have an impact on the price movement of the stock. So, to predict the future trend or price, the sentiment analysis on the available financial information is needed. The machine learning approach to extract sentiment from the financial information is more accurate than the Lexicon-Based sentiment analysis.

We have reviewed various off-the-shelf technologies available for sentiment analysis using machine learning. IBM Watson Natural Language Understanding is one such platform from IBM, which provides sentiment analysis on financial information with a great accuracy of 97% and can be trained further for specific domains to increase the accuracy. Also, this tool is commercially less expensive compared to its peers.

A future study is needed to understand the sentiment analysis using off-the-self technology such as IBM Watson Natural Language Understanding on the available financial information on both microeconomic and macroeconomic level of stocks in the Indian stock market and its impacts on the stock price movement in Indian stock market. New researchers can work further on this topic to create new forecasting models based on the latest machine learning algorithms and the sentiment analysis score from IBM Watson, along with the historical price of the stock. The prediction models will provide greater results than the previous models.

REFERENCES

1. Edwards R. D., Magee J., Bassetti W. H.C. Technical analysis of stock trends. New York, NY: AMACOM, a division of American Management Association; 2007. 840 p.
2. Nicholson C. Building wealth in the stock market: A proven investment plan for finding the best stocks and managing risk. Milton, Qld: John Wiley & Sons Australia, Ltd; 2009. 352 p.
3. Thomsett M. C. Practical trend analysis: Applying signals and indicators to improve trade timing. Boston, MA: Walter de Gruyter Inc.; 2019. 350 p.
4. Fama E. F. The behavior of stock-market prices. *The Journal of Business*. 1965;38(1):34–105. DOI: 10.1086/294743
5. Nagpal A., Jain M. Efficient market hypothesis in Indian stock markets: A re-examination of calendar anomalies. *Amity Global Business Review*. 2018;13(1):32–41.
6. Black F. Noise. *The Journal of Finance*. 1986;41(3):528–543. DOI: 10.1111/j.1540–6261.1986.tb04513.x
7. De Long J.B., Shleifer A., Summers L.H., Waldmann R.J. Noise trader risk in financial markets. *Journal of Political Economy*. 1990;98(4):703–738. URL: https://scholar.harvard.edu/files/shleifer/files/noise_trader_risk.pdf
8. Baker M., Wurgler J. Investor sentiment and the cross-section of stock returns. *The Journal of Finance*. 2006;61(4):1645–1680. DOI: 10.1111/j.1540–6261.2006.00885.x
9. Baker M., Wurgler J. Investor sentiment in the stock market. *Journal of Economic Perspectives*. 2007;21(2):129–151. DOI: 10.1257/jep.21.2.129
10. Kumari J., Mahakud J. Does investor sentiment predict the asset volatility? Evidence from emerging stock market India. *Journal of Behavioral and Experimental Finance*. 2015;8:25–39. DOI: 10.1016/j.jbef.2015.10.001
11. Kumari J., Mahakud J. Investor sentiment and stock market volatility: Evidence from India. *Journal of Asia-Pacific Business*. 2016;17(2):173–202. DOI: 10.1080/10599231.2016.1166024
12. Pang B., Lee L. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*. 2008;2(1–2):1–90. URL: https://www.cs.cornell.edu/home/llee/omsa/omsa.pdf
13. Misra P. An investigation of the macroeconomic factors affecting the Indian stock market. *Australasian Accounting Business & Finance Journal*. 2018;12(2):71–86. DOI: 10.14453/aabfj.v12i2.5

14. Kumar P., Gupta S. K., Sharma R. K. An empirical analysis of the relationship between FPI and Nifty returns. *IUP Journal of Applied Economics*. 2017;16(3):7–24.

15. Mann B. J. S., Babbar S. Stock price reaction around new product announcements: An event study. *IUP Journal of Management Research*. 2017;16(3):46–57.

16. Heston S. L., Sinha N. R. News vs. sentiment: Predicting stock returns from news stories. *Financial Analysts Journal*. 2017;73(3):67–83. DOI: 10.2469/faj.v73.n3.3

17. Allen D. E., McAleer M., Singh A. K. Daily market news sentiment and stock prices. *Applied Economics*. 2019;51(30):3212–3235. DOI: 10.1080/00036846.2018.1564115

18. Johnman M., Vanstone B. J., Gepp A. Predicting FTSE 100 returns and volatility using sentiment analysis. *Accounting & Finance*. 2018;58(S 1):253–274. DOI: 10.1111/acfi.12373

19. Chan S. W. K., Chong M. W. C. Sentiment analysis in financial texts. *Decision Support Systems*. 2017;94:53–64. DOI: 10.1016/j.dss.2016.10.006

20. Uhl M. W. Emotions matter: Sentiment and momentum in foreign exchange. *Journal of Behavioural Finance*. 2017;18(3):249–257. DOI: 10.1080/15427560.2017.1332061

21. Rani S., Singh J. Sentiment analysis: A survey. *International Journal for Research in Applied Science & Engineering Technology*. 2017;5(8)1957–1963. DOI: 10.22214/ijraset.2017.8276

22. Xing F. Z., Cambria E., Welsch R. E. Natural language based financial forecasting: A survey. *Artificial Intelligence Review*. 2018;50(1):49–73. DOI: 10.1007/s10462–017–9588–9

23. Yadav R., Kumar A., Kumar A. V. Event-based sentiment analysis on futures trading. *The Journal of Prediction Markets*. 2019;13(1):57–81. DOI: 10.5750/jpm.v13i1.1731

24. Shehu H. A., Tokat S., Sharif H., Uyaver S. Sentiment analysis of Turkish Twitter data. *AIP Conference Proceedings*. 2019;2183:080004. DOI: 10.1063/1.5136197

25. Alpaydin E. Introduction to machine learning. Cambridge, MA: The MIT Press; 2014. 537 p.

26. Carvalho A. Harris L. Off-the-shelf technologies for sentiment analysis of social media data: Two empirical studies. In: Americas conf. on information systems (AMCIS 2020). (August 15–17, 2020). Atlanta, GA: Association for Information Systems. 2020. URL: https://aisel.aisnet.org/amcis2020/social_computing/social_computing/6

27. de las Heras-Pedrosa C., Sánchez-Núñez P., Peláez J.I. Sentiment analysis and emotion understanding during the COVID-19 pandemic in Spain and its impact on digital ecosystems. *International Journal of Environmental Research and Public Health*. 2020;17(15):5542. DOI: 10.3390/ijerph17155542

28. Carvalho A., Xu J. Studies on the accuracy of ensembles of cloud-based technologies for sentiment analysis. In: Americas conf. on information systems (AMCIS 2021). (August 9–13, 2021). Atlanta, GA: Association for Information Systems. 2021:1462. URL: https://aisel.aisnet.org/amcis2021/art_intel_sem_tech_intelligent_systems/art_intel_sem_tech_intelligent_systems/12

29. Ince H., Trafalis T. B. A hybrid forecasting model for stock market prediction. *Economic Computation and Economic Cybernetics Studies and Research.* 2017;51(3):263–280. URL: http://www.eadr.ro/RePEc/cys/ecocyb_pdf/ecocyb3_2017p263–280.pdf

30. Cocianu C. L., Grigoryan H. Machine learning techniques for stock market prediction. A case study of Omv Petrom. *Economic Computation and Economic Cybernetics Studies and Research*. 2016;50(3):63–82. URL: https://www.researchgate.net/publication/308719462_Machine_learning_techniques_for_stock_market_prediction_Acase_study_of_OMV_Petrom

31. Moghaddam A. H., Moghaddam M. H., Esfandyari M. Stock market index prediction using artificial neural network. *Journal of Economics, Finance and Administrative Science*. 2016;21(41):89–93. DOI: 10.1016/j.jefas.2016.07.002

32. K.-S., Kim H. Performance of deep learning in prediction of stock market volatility. *Economic Computation and Economic Cybernetics Studies and Research*. 2019;53(2):77–92. DOI: 10.24818/18423264/53.2.19.05

33. Obthong M., Tantisantiwong N., Jeamwatthanachai W., Wills G. A survey on machine learning for stock price prediction: Algorithms ♦ and techniques. In: Proc. 2ⁿᵈ Int. conf. on finance, economics, management and IT business (FEMIB 2020). Vol. 1. Setúbal: Science and Technology Publications (SciTePress); 2020:63–71. DOI: 10.5220/0009340700630071

34. Rich H., Scott D., Franck B. Evaluating predictability of financial markets using New York Times sentiments and market data. 2017. URL: https://github.com/IBM/powerai-market-sentiment#readme

35. Wu G.G.-R., Hou T.C.-T., Lin J.-L. Can economic news predict Taiwan stock market returns? *Asia Pacific Management Review*. 2019;24(1):54–59. DOI: 10.1016/j.apmrv.2018.01.003

36. Rani N., Kaushal A., Shakir M.B. Social media and sentiment analysis of Nifty 50 Index. *Journal of Prediction Markets*. 2019;13(1):50–56. DOI: 10.5750/jpm.v13i1.1710

ABOUT THE AUTHORS

Anupa S. Dash — PhD Scholar, MIT College of Management, MIT Art, Design & Technology University, Pune, India
https://orcid.org/0000-0001-6359-2160
Corresponding author:
anups.dash@gmail.com

Ujjwal Mishra — PhD, Prof. of Finance, academician, MIT College of Management, MIT Art, Design & Technology University, Pune, India
https://orcid.org/0000-0002-3291-0143
ujjwalmmishra@gmail.com

Conflicts of Interest Statement: The authors have no conflicts of interest to declare.

The article was submitted on 02.09.2022; revised on 02.10.2022 and accepted for publication on 26.10.2022.
The authors read and approved the final version of the manuscript.