# Investment Strategies in Indian Stock Market Using Vader Algorithm Sentimental Analysis

**Dr. Geo Paul. K**
*Director, Monti International Institute of Management Studies, Perinthalmanna, Kerala, India*

**Dr. Nijo Varghese**
*Assistant Professor, Department of Management Studies, Naipunnya Business school, Pongam, Kerala, India*

**Dr. Nitha. K.P**
*Guest Faculty, Sree KeralaVarma College, Computer science department, Kerala, India*

**Dr. Suraj.E.S**
*Associate Professor, Department of Management Studies, Naipunnya Business school, Pongam, Kerala, India*

**Abstract**

Due to volatility and complexity of the stock market, it is challenging to predict and examine the critical factors that drives the stock market. Nevertheless, many people feel that investing in stock market is risky mainly because of complex factors determining the share prices in the market. Sometimes investors may lose more money from the stock market due to its risk. An understanding of the risks and various outside influences have an immense effect on the investment decisions. Market information, news, tips, market sentiments are an essential source for investing in stock market. Hence, the present study focused on evaluating the importance of sentiments of Sensex stocks driving stock prices in Indian stock market. This study suggests the stocks with positive sentiments, negative sentiments and its polarity score calculated using Vader algorithm for taking short term investment decisions to generate profit from stock market.

**Keywords:** *Vader Algorithm, Sensex Stocks, Sentiments, Polarity Score*

## 1. INTRODUCTION

Sentiment analysis refers to an automatic process of defining whether a piece of written information is having a positive, negative or neutral sentiment. It is also called opinion mining which is also mentioned as the discovery of subjective information such as opinions, attitudes, emotions, and feelings expressed by people in the form of blogs, tweets, and comments. According to surveys and studies conducted, 92% of the customers they trust the online reviews as much as their personal interests. But still, there are several challenges for accurate sentiment evaluation such as Struggle in training the computer to understand human sentences or linguistics, Gauging sentiments concerning several properties, Unavailability of standard scientific lexicons for parameters or features, to the domain being applied.

Share market with a vast number of features swaying the movement of the drift on changing scales and numerous layers is an impulsive system. Investor's sentiments affect the stock market to a great extent. Economic news heavily impacts the stock market as the economy goes up, and there by the stocks profitability also increases. This research investigates the how accurate the sentimental analysis can predict the market behaviour and also the movement of stock prices. Machine learning algorithms and different data extraction tools are capable enough to analyze and perform knowledge sighting at large scales in a short amount of time. Two methods for the prediction in the area of share market using sentimental analysis are:

- Qualitative Analysis

Newsfeeds concerning the stock market highly affect the market trend. The social networks, news, blogs, tweets about the stocks are all related and also, they together increase the unpredictability of the system. Prevailing study opinions that in case of emergency, stocks replicate each other and lead to market crashes. Today, Twitter has become the most dependable and firmest way of consuming media. Sentiments about a company can be very easily identified with these combined resources of the news feed and Twitter feed. The best tools for this kind of analysis are text mining and sentiment analysis.

- Quantitative Analysis

Past data is today readily accessible for most of the scenarios. By means of this dataset, several machine learning models can be applied to give exact outcomes for forthcoming investments. With the most reflective variables these models can be trained for separate stocks. These models can also be competent to work in diverse scenarios and overall market movement.

The commonly used approaches concentrates on fundamental and technical analysis to predict the market at a large scale, which seldom interprets to low-level distinct stock forecast. But individual stocks instead of creating an impact to their specific performance, contributes to the entire market movement. Thus, focusing on individual stocks to predict market movement is a much more adaptable approach. With technology progressing at such a rapid pace and availability of computing power a inclusive system can attempt towards to accurately predict the market trend and reap valuable economic earnings. Current studies proves that the modern approach outperforms the traditional approach and can give most accurate results.

Identifying whether the expressed sentiment is positive, negative or neutral for a given writing with extreme accuracy is the most critical task in the sentimental analysis. In this study, the sentiment polarity of the news headlines concerning each stock for 10 financial years (2010-11 to 2020-21) is calculated and its average is considered as sentiment value i.e., the polarity of the stock to next module. This approach is used to predict the buy or sell signal for the investors, which helps them in deciding while investing in the stock market. News headlines are collected from financial news sites moneycotrol.com. For the analysis, only the text data of each news or article is used. Twitter data collected through the public API. The data with the hashtags of Sensex stocks alone is retrieved. The VADER algorithm is a powerful lexicon-based sentiment analyzer algorithm that uses semantic data from several corpora to analyze the sentiment of each statement. This algorithm does not require a training data set which is one of the advantages. Study mainly focus on identifying the short term investment strategies for the Sensex stocks based on the sentiments prevailing in the market.

## 2. REVIEW OF LITERATURE

Geetha et al. (2017) read the slant Analysis for successful securities exchange forecast. This paper plan and execute the prescient framework for directing financial exchange speculation.

This investigation utilized the blend of both Sensex focuses and truly Simple Syndication (RSS) channels for compelling expectation. Assessment investigation of news sources impacts stock costs. This examination utilized a calculation for conclusion investigation through the connection between the financial exchange esteems and assessments in RSS news sources and a prepared model is utilized for the expectation of stock costs. Sahar Sohangir et al. (2018) utilized Big Data procedures and profound learning models for nostalgic examination. Profound Learning models can be utilized to improve the exhibition of assessment investigation for stock tweets. Consequences of the investigation show that the profound learning model can be utilized successfully for monetary estimation examination and a convolutional neural organization is the best model to anticipate the feeling of creators in the Stock Tweets dataset. It was demonstrated that profound learning perhaps the best model for supposition investigation.

Zhaoyue Wang et al. (2018) utilized bimodal calculation with Data-Divider to Predict Stock Index. Existing calculations, including Backpropagation and numerous current calculations, don't give supportive expectation results to stock financial backers. Stock information for extensive stretches is excessively mind boggling and incorporate such a large number of modes, yet "old" BP neural organizations are as yet potential on the lookout. Sigo et al. (2018) utilized large information examination use of counterfeit neural organization in estimating stock value drifts in India. This paper investigations the nonlinear example in the offer value development of the most unpredictable three stocks as far as market capitalization, recorded in Bombay Stock Exchange (BSE) in India, like Reliance Industries Limited, Tata Consultancy Service Limited and HDFC Bank Limited, utilizing Artificial Neural Network for the examination time frame from 2008 to 2017. This investigation anticipated the upsides of huge market promoted organizations, specifically RIL, TCS and HDFC Bank, that are recorded in the Bombay Stock Exchange. Data acquired were pre-prepared productively utilizing AI strategy like the counterfeit neural organization will give more exact forecast and the financial backers could guarantee procuring capital appreciation, for their corporate shares.

Siddikee et al. (2018) contemplated the effect of every day profit on math and logarithmic return analysed the measurable strength of Model-1 created in the examination. Study uncovered that the incorporation of day-by-day profits fundamentally expanded the everyday and month to month number-crunching and logarithmic returns of the protections. Also, after the incorporation of the day by day profit the Value at Risk (VaR) of the day-by-day logarithmic return decays forcefully approves Model 1 for figuring the day-by-day logarithmic return. Wanzala et al. (2018) assessed the market instantaneousness by Coefficient of Elasticity of Trading. This technique is utilized to empower financial backers to smooth utilization or to change the danger return skylines of their portfolios. This investigation propounds another proportion of market instantaneousness (that is, CET3) and goes further to work with correlation with other two proportions of promptness (that is, CET1 and CET2) through OLS relapse examination utilizing information from 2001 to 2016 from NSE and KNBS in Kenya. Studies have demonstrated that macroeconomic factors didn't substantially influence the connection between market quickness and monetary development. Relapse esteems and standard blunders exhibit that CET3 is a vigorous proportion of market quickness than CET1 and CET2.

Mudinas et al. (2018) utilizing Sentiment Analysis for market pattern forecast in this examination. This examination explores the capability of utilizing opinion perspectives (positive versus negative) and furthermore supposition feelings (happiness, misery, and so forth) extricated from monetary news or tweets to help foresee stock value developments. It was seen that overall assumption perspectives don't appear to Granger-cause stock value changes and explicit events assessment feelings do appear to Granger-cause stock value changes. Thus, the example shown isn't all inclusive and should be taken a gander at dependent upon the situation. Incorporating feeling feelings as extra highlights into the AI based market pattern expectation model could improve its precision.

Shantanu Pacharkar et al. (2018) predicting financial exchange speculation Using conclusion investigation shown that a solid relationship exists between rising/fall in stock costs of an organization to the popular sentiments or feelings. A notion analyser has been created in the examination to pass judgment on the sort of notion present in the survey. The surveys are characterized into three classifications: positive, negative and impartial. An arrangement calculation would order each survey as being either certain or negative utilizing a model is made from a bunch of preparing information Instead of putting resources into an organization whose end costs are high. On the off chance that the wistful score is high and positive, there are high possibilities at its stock costs to go up later on. Gupta et al. (2018) utilized nostalgic investigation on news information for financial exchange expectation. This investigation gathered the news, latest things, and memorable cost of the stock. It was tracked down that the technique for dissecting social information with noteworthy cost will give us great outcomes like 80% exactness. The task considers both news and noteworthy cost to anticipate the financial exchange development precisely in the securities exchange.

Bhavya Kaushik et al. (2017) designed to examine web-based media (SM) and what it means for their stock costs. The motivation behind the paper is to know the connection between SM use and the organization's stock cost. This paper relates NSE stock value patterns of each firm with its SM use and SM prevalence, utilizing apparatuses like connection, relapse, and ANOVA. It was discovered that there is no agreeably reasonable or numerically expressible connection between the stock costs of firms and their separate SM exercises. Umamaheswari et al. (2018) considered the utilization of Sentimental Analysis and Machine learning calculations, assisted financial backers with putting their cash in shares This examination furnish the client with two choices to anticipate a similar utilizing either nostalgic investigation or through AI calculations.

Tahir M. Nasir et al. (2018) investigate the connection between legislative issues related slant and FTSE 100 developments. This examination was led on a short-window occasion investigation of a UK-based political occasion to know the proof of a relationship between the overall state of mind of the general population and speculation conduct for the time being. It was apparent that causation between open conclusion and the financial exchange development's relationship exists. Thus, these outcomes show guarantee for utilizing slant investigation on Twitter information for gauging market developments. Ghaith Abdulsattar A et al. (2017) studied the job of estimation examination, Twitter information, and highlights in foreseeing securities exchange returns. It was tracked down that the opinions and feelings of individuals can likewise help in anticipating securities exchange returns. Highlights like spatial and worldly can likewise influence the securities exchange returns. The spatial component is a geological division, it can either be various feelings of various individuals from various topographical districts, or they can be other securities exchanges that can influence the home financial exchange by any connection.

## 3. OBJECTIVES OF THE STUDY

a) To conduct sentimental analysis for Sensex stocks using Vader Algorithm.

b) To identify the stocks with positive, negative and neutral sentiments for short term investment decision.

c) To suggest the short term investment strategies based on the polarity score using sentimental analysis.

## 4. SOURCES OF DATA

An analytical study with the secondary sources of data including the yearly reports, official websites of BSE (www.bseindia.com), NSE (www.nseindia.com), Confederation of Indian Industries (CII) library, Centre for Monitoring Indian Economy (CMIE) Prowess and annual report library services like Ebsco Services is considered. Additional resources like, working papers, research papers, and economic dailies are also referred for this purpose. Financial data from 2010-11 to 2020-21(10 years) is collected for the BSE stocks. Table 1.1 exhibits the list of stocks in the BSE stocks.

4.1. Sentimental analysis-Methodology

1. Conduct sentimental analysis to know the opinion, emotions expressed by people about each stock.

2. Accuracy of sentimental analysis will be measured by using yearly returns of each stock during the next period to be compared with buy/sell recommendation.

**Table 1.1 List of Sensex Stocks**

| **Sl. No** | **Name of the Stock** |
|---|---|
| 1 | Axis Bank |
| 2 | Asian Paints |
| 3 | Bajaj Auto |
| 4 | Bajaj Finance |
| 5 | Bharathi Airtel |
| 6 | Coal India |
| 7 | HCL Tech |
| 8 | HDFC |
| 9 | HDFC Bank |
| 10 | Hero Motors |
| 11 | Hindustan |
| 12 | ICICI Bank |
| 13 | IndusInd Bank |
| 14 | ITC |
| 15 | Kotak Bank |
| 16 | Larsen & Toubro (L&T) |
| 17 | Maruti |
| 18 | Mahindra & Mahindra (M&M) |
| 19 | National Thermal Power Corporation Ltd (NTPC) |
| 20 | Oil and Natural Gas Corporation Limited (ONGC) |
| 21 | Power Grid |
| 22 | Reliance |
| 23 | State Bank of India |
| 24 | Sun Pharma |
| 25 | Tata DVR |
| 26 | Tata Motors |
| 27 | Tata Steel |
| 28 | Tata Consultancy Services (TCS) |
| 29 | Vedanta |
| 30 | Yes Bank |

4.2. Softwares

To analyze the data, R Software 4.0.2 is used. To develop a model using sentimental indicators, R python, Cocalc is used.

4.3. Methodology for calculating the Polarity of the text using the VADER Algorithm

Figure 5.1 gives a detailed view of the different steps used in calculating the polarity of the statements using the VADER Algorithm. The steps used in this module are as follows:

a. Data Collection- The data is collected using the pandas' data reader to scrape news headlines from moneycontrol.com. The developer mode API of twitter.com is used to extract data with the hash tags of the Sensex stocks, from Twitter.

b. Calculating polarity of the headlines using the 5 Heuristics of the VADER Algorithm- Each news headline and the Twitter data are broken down into sentences. Further, the sentences re broken down into words to find the VADER scores based on the positive or the negative score as well as the intensity of emotions considering the following 5 heuristics:

•Punctuations: They increase the magnitude of intensity. For example, the sentence,” The price is high !!!” with punctation exclamations increases the intensity of the sentences when compared to the sentence,” The price is high”.

• Capitalization: All letters of a word in caps usually emphasize the sentiment-relevant word. For example, consider the statements, “The price is HIGH” and “The price is high". The word HIGH in the first sentence highlights the relevance of the high price of the stock when compared to the "high" in the second statement.

• Degree Modifiers: The words like extremely, very, pretty, fairly are considered to impact the sentiment of sentences. For example, “The price is extremely high” intensifies the sentiment of the statement when equated to “The Price is high”

• Conjunctions: The use of constructive conjunctions like ‘but', 'nor', 'so' , 'or' , 'yet, signals a shift in sentiment polarity, giving dominance to one part of the sentences. For example, “The price is high, but it is affordable”, the contrastive conjunction “but” used in the above sentence signals a shift in sentiment polarity, giving dominance to the latter half.

• Polarity Negation: The negations used in the statements tend to flip the polarity/sentiments of the text. A contiguous sequence of 3 words, preceding a sentiment-laden lexical feature and a negation has more impact on the polarity of the statement. Consider the statement” The price isn't really that high”, the actual polarity of the statement is positive, but if the 3 contiguous sequence of the word “really that high” along with the negation “isn’t” is not considered, then it might be categorized as a negative statement.

c. Summing and Calculating the final VADER Scores

After calculating the individual scores of each word of the sentences the final compound scores are calculated, as the average of the individual scores, and the polarity is determined by mapping them to the threshold value.

d. Mapping the final compound scores to the threshold, categorizing them as positive, negative, and neutral

In this step the aim is to map the compound scores to the threshold and find the polarity of the text. The following rules are applied to identify the polarity from the compound scores.

• If the Compound Score>0.6 then the text highlights a positive sentiment,

• If the Compound Score is between -0.5 and 0.54 then orientation of the sentiment is considered to be neutral,

• If the Compound Score < -0.5 then the text gives a negative sentiment.

Sentiment analysis for the extracted news headlines and tweets are performed using the VADER algorithm. From this, positive, negative, and neutral individual scores, and the compound sentiment scores are obtained for each statement for a given stock for the 10 financial years. Each of these sentiment scores is in the range of -1 to 1. If the VADER score is greater than 0.6 it is considered being the most positive sentiment and less than -0.5 is the most negative sentiment.

If the Compound Score is between -0.5 and 0.54 then it is considered to be Neutral. The compound score is characteristic of all these three lexicons and the average of the compound sentiment score for each headline and tweets of a year is averaged to obtain a sentiment score. VADER score can be calculated to know about the market sentimental score from the news also like this. If the VADER score calculated lies in between 0 and 1 indicates that the market will be bullish on the given day. If the score is between -1 and 0 suggests that the fiscal world has a negative view on the market, will be bearish for the day. If the sentiment score obtained is 0, then there is neutral sentiment for the market on that given day.

## 5. ANALYSIS

Study show the descriptive statistics of polarity score for 30 Sensex stocks. High Mean value of polarity score for Asian paints, Bajaj Finance, Bharathi Airtel, ITC, Hindustan Unilever, Maruti Suzuki and Tata steel show that news and prospects of the company remained positive for the last ten years. The highest mean shows the confidence of investors to remain with the same company for long-term. Minimum of polarity score at a very low value for Mahindra & Mahindra (MM) indicates that the news and sentiments of the company remained negative for that specific year. It indicates the massive selling of the stock for that year. It is not advisable to buy the stocks with extreme minimum value of the polarity score

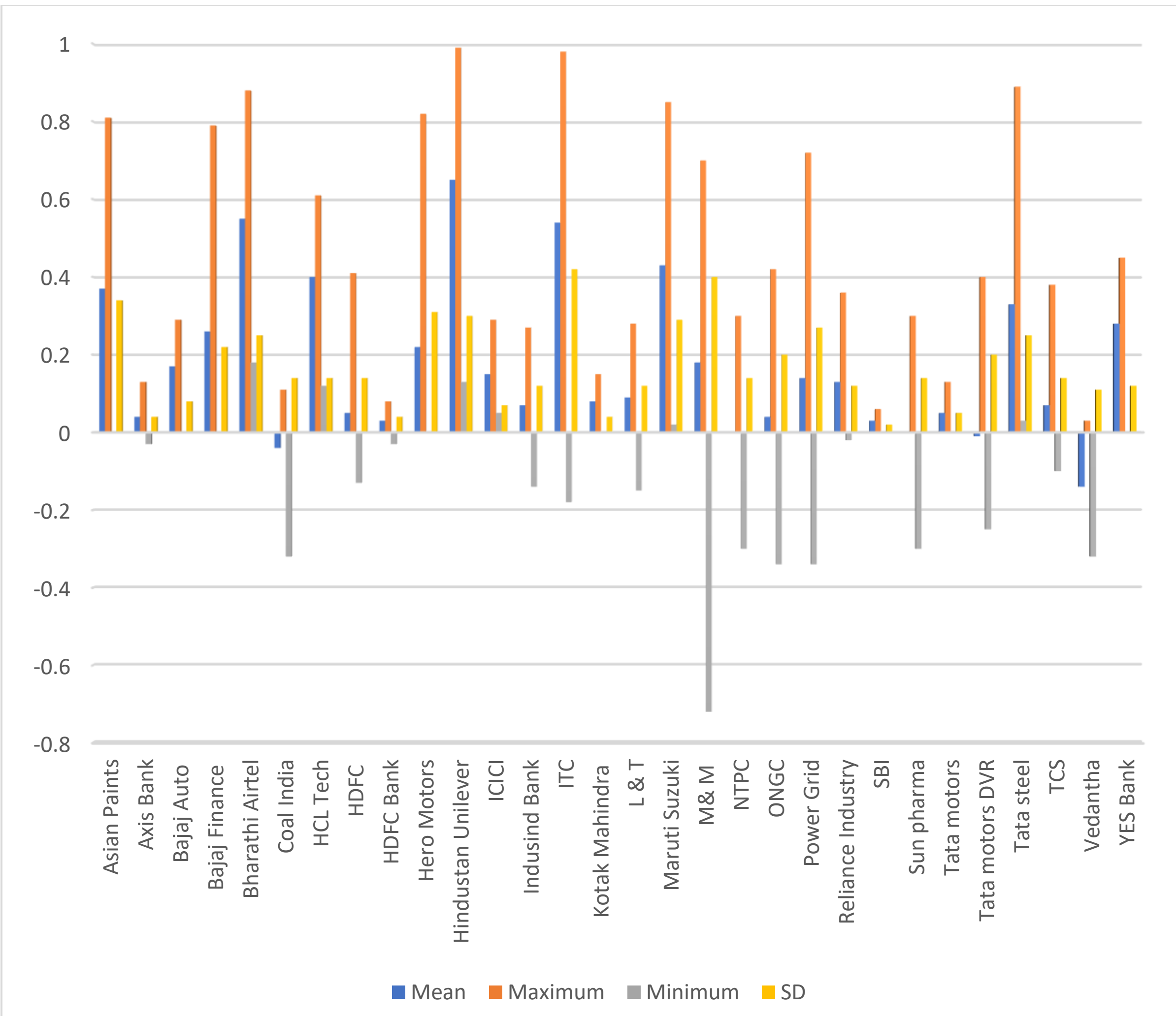
**Table 1.2-Descriptive Statistics of the Polarity Score of the Sensex Stocks from 2010-11 To 2020-21**

| Name of stock | Minimum | Maximum | Mean | SD |
|---|---|---|---|---|
| Asian Paints | 0.00 | 0.81 | 0.37 | 0.34 |
| Axis Bank | -0.03 | 0.13 | 0.04 | 0.04 |
| Bajaj Auto | 0.00 | 0.29 | 0.17 | 0.08 |
| Bajaj Finance | 0.00 | 0.79 | 0.26 | 0.22 |
| Bharathi Airtel | 0.18 | 0.88 | 0.55 | 0.25 |
| Coal India | -0.32 | 0.11 | -0.04 | 0.14 |
| HCL Tech | 0.12 | 0.61 | 0.40 | 0.14 |
| HDFC | -0.13 | 0.41 | 0.05 | 0.14 |
| HDFC Bank | -0.03 | 0.08 | 0.03 | 0.04 |
| Hero Motors | 0.00 | 0.82 | 0.22 | 0.31 |
| Hindustan Unilever | 0.13 | 0.99 | 0.65 | 0.30 |
| ICICI | 0.05 | 0.29 | 0.15 | 0.07 |
| IndusInd Bank | -0.14 | 0.27 | 0.07 | 0.12 |
| ITC | -0.18 | 0.98 | 0.54 | 0.42 |
| Kotak Mahindra | 0.00 | 0.15 | 0.08 | 0.04 |
| Larsen & Toubro | -0.15 | 0.28 | 0.09 | 0.12 |
| Maruti Suzuki | 0.02 | 0.85 | 0.43 | 0.29 |
| M& M | -0.72 | 0.70 | 0.18 | 0.40 |
| NTPC | -0.30 | 0.30 | 0.00 | 0.14 |
| ONGC | -0.34 | 0.42 | 0.04 | 0.20 |
| Power Grid | -0.34 | 0.72 | 0.14 | 0.27 |
| Reliance Industry | -0.02 | 0.36 | 0.13 | 0.12 |
| SBI | 0.00 | 0.06 | 0.03 | 0.02 |
| Sun pharma | -0.30 | 0.30 | 0.00 | 0.14 |
| Tata motors | 0.00 | 0.13 | 0.05 | 0.05 |
| Tata motors DVR | -0.25 | 0.40 | -0.01 | 0.20 |
| Tata steel | 0.03 | 0.89 | 0.33 | 0.25 |

| | | | | |
|---|---|---|---|---|
| TCS | -0.10 | 0.38 | 0.07 | 0.14 |
| Vedanta | -0.32 | 0.03 | -0.14 | 0.11 |
| YES Bank | 0.00 | 0.45 | 0.28 | 0.12 |
| *Note: Results computed using R 4.0.2* | | | | |

Maximum of polarity score at very high value for ITC, HUL indicates that the news and sentiments for the same company remained positive for that year under consideration. It is advisable to buy the stocks with high maximum value of the polarity score. The extreme value of the standard deviation of polarity score shows the variability of sentiments for M&M which can be considered as outliers of the sample. Even though ITC has high value of standard deviation it is not considered as an outlier since it does not have low value for minimum polarity score.

**Figure 1.1. Polarity Score of Sensex Stocks from 2010-11 To 2020-21**




VADER scores considering the 5 heuristics such as Punctuations, Capitalization, Degree Modifiers, Conjunctions, Polarity Negation to calculate the individual scores of each word of the sentences the and calculating the final score based on its average.Rules for determining the category of sentiment given below

Rule 1: If the Compound Score>0.6 then the text highlights a positive sentiment,

Rule 2: If the Compound Score between -0.5 and 0.6 the it is considered to be Neutral

Rule 3: If the Compound Score < -0.5 then the text gives a negative sentiment.

VADER algorithm is a powerful lexicon-based sentiment analyzer used to analyze the sentiment of each sentence. 558 % of sentiments are positive, 11 % are negative and 31 % are neutral. There is no similarity between percentage of negative and positive sentiments. It suggests that Sensex stocks carry more positive sentiments in the market for last ten years.

**Table1.3: Mean of Positive Polarity, Neutral Polarity and Negative Polarity**

| Name of stock | Mean of Positive Polarity | Mean of Negative Polarity | Mean of Neutral Polarity |
|---|---|---|---|
| Asian Paints | 0.52 | * | 0.03 |
| Axis Bank | 0.09 | * | 0.02 |
| Bajaj Auto | 0.18 | * | 0.15 |
| Bajaj Finance | 0.26 | * | 0.00 |
| Bharathi Airtel | 0.59 | * | * |
| Coal India | 0.09 | -0.24 | 0.02 |
| HCL Tech | 0.16 | * | * |
| HDFC | 0.20 | -0.13 | 0.00 |
| HDFC Bank | 0.08 | * | 0.01 |
| Hero Motors | 0.64 | * | 0.00 |
| HUL | 0.66 | * | * |
| ICICI | 0.17 | * | 0.16 |
| IndusInd Bank | 0.16 | -0.06 | 0.01 |
| ITC | 0.70 | -0.18 | 0.02 |
| Kotak Mahindra | 0.09 | * | -0.01 |
| Larsen & Toubro | 0.15 | -0.15 | 0.00 |
| Maruti Suzuki | 0.51 | * | 0.02 |
| Mahindra & Mahindra | 0.39 | -0.72 | 0.00 |
| NTPC | 0.30 | -0.30 | 0.00 |
| ONGC | 0.36 | -0.34 | 0.03 |
| Power Grid | 0.26 | -0.18 | 0.17 |
| Reliance | 0.17 | -0.02 | 0.03 |
| SBI | 0.06 | * | 0.02 |
| Sun pharma | 0.30 | -0.30 | 0.00 |
| Tata motors | 0.05 | * | * |
| Tata DVR | 0.26 | -0.14 | -0.05 |
| Tata steel | 0.36 | * | 0.03 |
| TCS | 0.26 | -0.01 | 0.03 |
| Vedanta | * | -0.19 | 0.03 |
| Yes | 0.32 | -0.01 | * |
| *Note: Results Obtained by using R 4.0.2, * Data not available* | | | |

In the Sentimental Analysis, the lexicon identified are primarily- positive, negative, and neutral. Sentiment analysis is applied in study to forecast “bullish” and “bearish” stocks from the data that revealed sentiment about the stock tips or tweets, collected from the media. The sentiments/polarity is calculated for the news headlines and tweets. The polarity identified, depicts the buy/sell signals for the stocks. The financial news doesn't follow a structured format. Sentimental analysis of individual stocks traded in the Indian share market is difficult to apply with greatest accuracy due to the lack of this structured format.Table 1.3 focused on the mean of positive, negative, and neutral score of the sentiment's polarity calculated on the financial news taken from the market. Overall emotions, both at the level of groups of traders and investors, are influencing the behaviour of financial markets. The strategies applied by portfolio managers with several assumptions on portfolio management is based on sentimental analysis. They can use stocks with high mean of polarity score for constructing the portfolio for providing returns to the investors. Asian paints, Tata steel, Bharathi airtel, and Hindustan Unilever have high polarity positive score which tells the positive sentiments, news and prospects of the company. The highest mean of positive sentiments shows the confidence of investors to remain with these companies. Extreme value of standard deviation of polarity score shows the variability of sentiments for Mahindra & Mahindra and ITC. Mean value of negative polarity score for Coal India, NTPC, Sun pharma, Vedanta indicates the negative sentiments of these stocks in the market.

**Fig 1.2 Mean of Positive Polarity Score Of The Sensex Stocks**

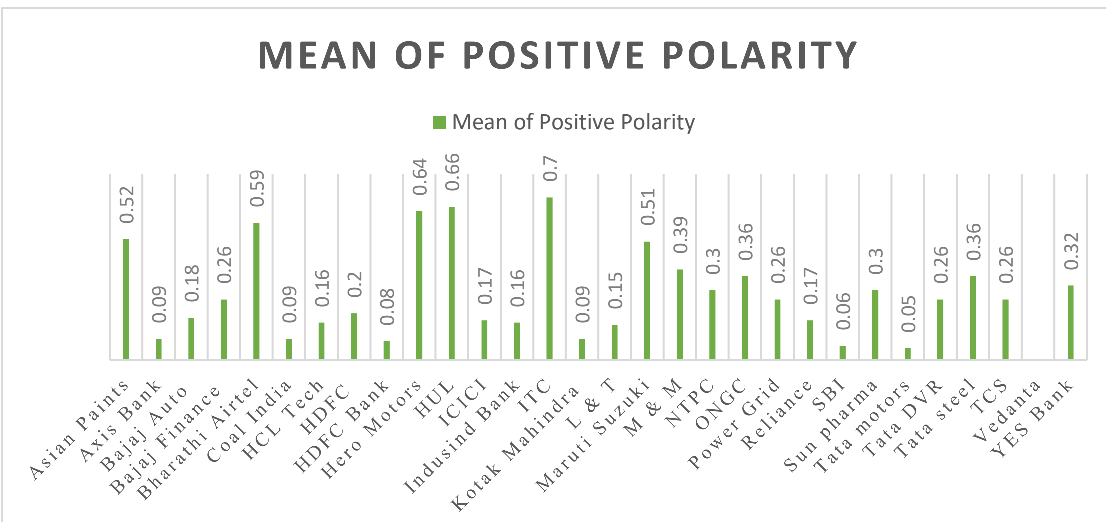



Fig 1.2 shows the mean positive polarity score of the Sensex stocks. The highest mean of positive polarity score is for Asian paints, Bharathi Airtel, Hero Motors, HUL, ITC and Maruti Suzuki. Sentiments of these stocks are highly positive, which influences the price movement in favour of the investors.

**Fig 1.3 Mean of Negative Polarity score of the Sensex Stocks**

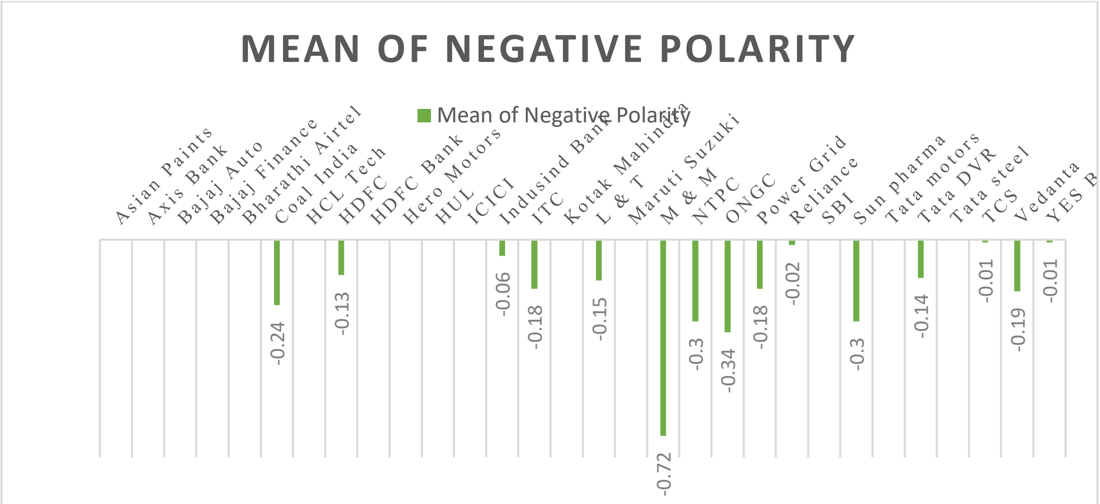



Fig 1.3 shows the Mean of the Negative Polarity score of the Sensex stocks. The mean value of negative polarity score for Mahindra & Mahindra, indicates the negative sentiments of the stock in the market. There is an adverse effect of sentiments on price movement for the stocks which is not in favour for investors.

**Fig 1.4 Mean Of Neutral Polarity Score of the Sensex Stocks**

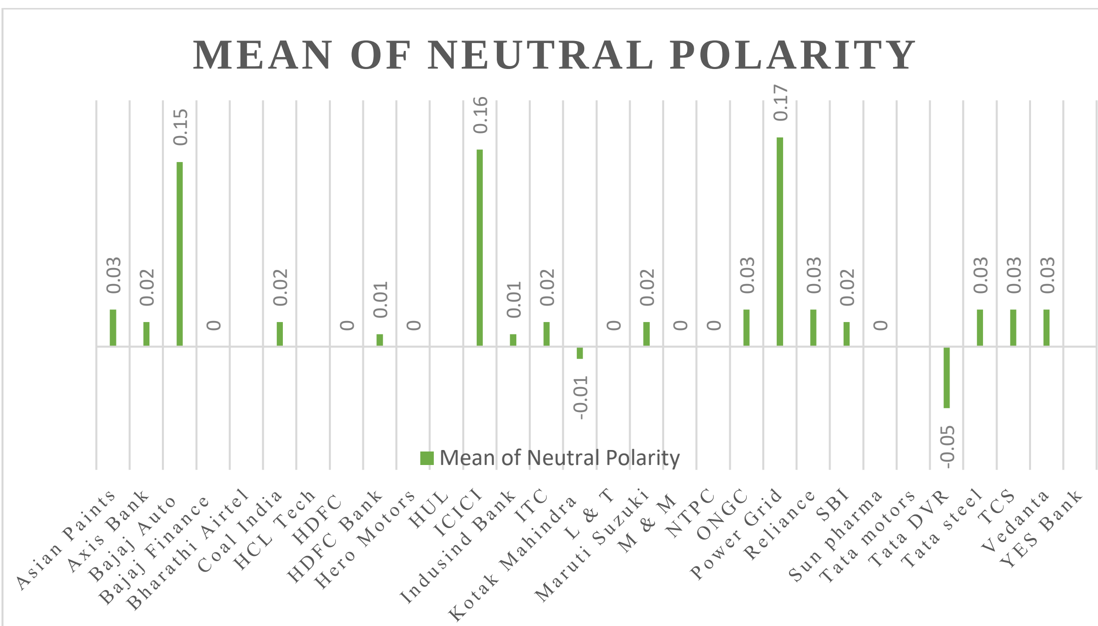



Fig 1.4 shows the Mean of the Neutral Polarity score of the Sensex stocks. The mean value of neutral polarity score for Bajaj Auto, ICICI, Power Grid indicates the neutral sentiments of these stocks in the market. There is no significant effect of sentiments on the price movement for these stocks.

## 6. CONCLUSION

In the sentimental analysis, when positive news appeared, it is hypothesized that the stock price of the company will increase and vice versa thereafter. It will take a lot of time to read, understand the news, its significance, and then make an investment decision. This average polarity score calculated for the last ten years can be used as a reference for the traders to trade in the above Sensex stocks. Investors can use this method to make investment decisions more efficiently. But there are some shortfalls of sentimental analysis in the study. Computers perform sentiment analysis, which provides the sentiment of the text. Several reviews investigate the influence of neutral word removal on the performance of sentiment analysis. However, neutral words are not as important for sentiment analysis as emotion words, which indicate an emotion of strong polarity. Sentiment score of the news headlines regarding each stock for 10 financial years calculated and its average is considered as sentiment value i.e., the polarity of the stock to classification module for Buying and selling the stocks. VADER scores considering the 5 heuristics such as Punctuations, Capitalization, Degree Modifiers, Conjunctions, Polarity Negation to calculate the individual scores of each word of the sentences the and calculating the final score based on its average. 558 % of sentiments are positive, 11 % are negative and 31 % are neutral. There is no similarity between percentage of negative and positive sentiments. It suggests that Sensex stocks carry more positive sentiments in the market for last ten years..Asian paints, Tata steel, Bharathi airtel, and Hindustan Unilever have high polarity positive score which tells the positive sentiments, news and prospects of the company. The highest mean of positive sentiments shows the confidence of investors to remain with these companies.

## References


Geetha, Shri Bharathi, Angelina. “Sentiment Analysis for Effective Stock Market Prediction.” International Journal of Intelligent Engineering and Systems (2017): 146-154.

Sahar Sohangir, Dingding Wang, Anna Pomeranets and Taghi M. Khoshgoftaar. “Big Data: Deep Learning for Financial Sentiment Analysis.” Journal of big data (2018): 1-25.

Zhaoyue Wang, Jinsong Hu , and Yongjie Wu. “A Bimodel Algorithm with Data-Divider to Predict Stock Index.” Hindawi-Mathematical Problems in Engineering (2018): 1-14. 137.

Sigo, Marxia Oli. “Big Data Analytics-Application of Artificial Neural Network In Forecasting Stock Price Trends In India .” Academy of Accounting and Financial Studies Journal (2018): 1-13.

Siddikee, Md. Noman. “Effect of Daily Dividend on Arithmetic and Logarithmic Return.” The Journal of Finance and Data Science (2018): 247-272.

Wanzala, Richard Wamalwa. “Estimation of market immediacy by Coefficient of Elasticity of Trading three approaches.” The Journal of Finance and Data Science (2018): 139-156.

Mudinas, Andrius. “Market Trend Prediction using Sentiment Analysis:Lessons Learned and Paths Forward.” WISDOM’18, August 2018, London, UK (2018): 1-14

Shantanu Pacharkar, Pavan Kulkarni. “Predicting Stock Market Investment Using Sentiment Analysis.” International Journal of Advanced Research in Computer and Communication Engineering (2018): 109-114.

Gupta,Juhi. “Sentimental analysis on News Data for Stock Market Prediction.” International Journal of Management and Applied Science (2018): 84-86.

Bhavya Kaushik, Harshit Hemani. “Social media usage vs. stock prices: an analysis of

Indian firms." Procedia Computer Science (2017): 323-330.

Umamaheswari, K.M. "Stock Market Predictor and Analyser using Sentimental Analysis and Machine Learning Algorithms." International Journal of Pure and Applied Mathematics (2018): 15395-15405.

Tahir M. Nasir, Man Yeung. "Twitter as a tool for forecasting stock market movements: A short-window event study." The Journal of Finance and Data Science (2018).

Ghaith Abdulsattar A.Jabbar Alkubaisi, Siti Sakira Kamaruddin. "A Systematic Review on The Relationship between Stock Market Prediction Model Using Sentiment Analysis on Twitter based on Machine Learning Method and Features. ." Journal Of Theoretical and Applied Information Technology (2017): 6924-6933.

Das, Debashish. "Hybrid Clustering-GWO-NARX neural network technique in predicting stock price." IOP Conf. Series: Journal of Physics: Conf. Series (2017): 1-14. 36.

Das, Sushree. "Real time sentiment analysis of twitter streaming data for stock prediction." Procedia Computer Science (2018): 956-964.