ISSN:0975-9646

Spandan Ghose Chowdhury et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (3) , 2014, 3595-3604

News Analytics and Sentiment Analysis to Predict Stock Price Trends

Spandan Ghose Chowdhury , Soham Routh , Satyajit Chakrabarti

Department of Computer Science and Engineering

Institute of Engineering & Management
Kolkata, India

***Abstract*— Business news carries varied information of different companies. But in this rapidly moving world the number of news sources present is uncountable, and it's humanly impossible to read and find all relevant information in the form of news to draw a conclusion timely to make an investment plan that returns maximum profit. In this paper, we have proposed a predictive model to predict sentiment around stock price. First the relevant real time news headlines and press-releases have been filtered from the large set of business news sources, and then they have been analyzed to predict the sentiment around companies. In order to find correlation between sentiment predicted from news and original stock price and to test efficient market hypothesis, we plot the sentiments of 15 odd companies over a period of 4 weeks. Our result shows an average accuracy score for identifying correct sentiment of around 70.1%. We also have plotted the errors of prediction for different companies which have brought out the RMSE and MAE of 30.3% and 30.04% respectively and an enhanced F_1 factor of 78.1%. The comparison between positive sentiment curve and stock price trends reveals 67% co-relation between them, which indicates towards existence of a semi-strong to strong efficient market hypothesis.**

***Keywords*— stock price trends, prediction model, knowledge discovery, sentiment analysis, market trends, news analytics, efficient market hypothesis.**

I. Introduction

In this closely connected world, we have access to all sorts of information. Hence, it's easier for us to make informed choices about any topic. Our decisions constantly get influenced by the kind of news we come across on day to day basis. The sentiment towards a particular entity becomes the driving force to take a proper decision.

There is a strong yet complicated relation between the market and the information available in the form of news. The arrival of news at every moment changes the perception or sentiment towards a particular company or their adopted business strategies. These days due to the bliss of internet, the traders, and investors have constant access to the updated news, and the news constantly mould their sentiments and influences their decision to invest in a particular company. The news from standard and authentic sources helps them to rebalance their assets judiciously. This sentiment not only becomes a deciding factor while rebalancing funds, but also aids in the decision of risk controls, and hence the stock market reflects the sentiment.

Major news can significantly impact on traders' sentiments [1] [2] [3] [4] [5] [6] yielding to a major change in the entire investment plan. Every investment, small or big, has significant impact on the growth of the company. When a trader having an enhanced positive sentiment finds a company worth investing, the company's growth also gets boosted up. On the other hand, the trader having a negative perception about a particular doesn't feel safe investing in a company, withdraws investments in order to avoid loss, and as a result of which the stock price of the company goes down drastically. [7] [8]

Considering how news impacts markets, Barber and Odean [9] note ''significant news will often affect investors' beliefs and portfolio goals heterogeneously, resulting in more investors trading than is usual'' (high trading volume). It is well known that volume increases on days with information releases [10]. Important news frequently results in large positive or negative returns. Ryan and Taffler (2002) find for large firms a significant portion (65%) of large price changes and volume movements can be linked to publicly available news releases. Sometimes investors may find it difficult to interpret news resulting in high trading volume without significant price change. [11] [12]

News has always been an important source of information to build perception of market investments. As the volumes of news and news sources are increasing rapidly, it's becoming impossible for an investor or even a group of investors to find out relevant news from the big chunk of news available. But, it's important to make a rational choice of investing timely in order to make maximum profit out of the investment plan. And having this limitation, computation comes into the place which automatically extracts news from all possible news sources, filter and aggregate relevant ones, analyse them in order to give them real time sentiments.

In this paper we have collected, classified and analysed relevant real time news from widely accepted and trusted news sources available in the public domain to forecast the perception around a particular company which helps the investors and traders to come up with informed, rational investment plan timely which ensures maximized profit. In this study we also have tracked the sentiments of the companies over a period of 4 weeks in order to find out co-relation between the moods predicted from the relevant news and the original stock price curve's ups and downs.

The main contributions of the paper are as follows:

- Implementation of an efficient prediction model that scores sentiments from all relevant real time news available in over 15 news sources present in public domain.

- This model classifies all of the relevant real time news from the large database of news.
- This forecasts real time news sentiment that reflects stock price movement trends.
- Compares and finds out a strong mean co-relation of 67% between positive sentiment movement and original stock price curve of 15 companies over a period of a month, and hence proves the test of EMH.
- The accuracy level of this prediction model is 70.1%. The prediction model has Mathew's Correlation Coefficient of 0.4.

The rest of the paper is organized as follows. Section II outlines the existing work in this field of sentiment analysis and concept of Efficient Market Hypothesis. Section III proposes and describes our predictive model and further plots sentiment trend predicted from the model and compares with original stock price movements. In next section i.e. section IV we evaluate the prediction model using different accuracy measurement techniques. In this section, we also compare and assess the co-relation of the sentiment trends of the proposed model with the original stock price movement. Finally in Section V, we conclude our paper and discuss the future work related to it.

II. Literature Review

A. Sentiment Analysis

Several systems have been built which attempt to quantify opinion from product reviews. Pang, Lee and Vaithyanathan [13] perform sentiment analysis of movie reviews. Their results show that the machine learning techniques perform better than simple counting methods. In another paper [14], they identified which sentences in a review are of subjective character to improve sentiment analysis. They did not make this distinction in their system, because they felt that both fact and opinion contribute to the public sentiment about news entities. [15] [16] [17]

Wilson, Wiebe, Hoffman [18] [19] presented a new approach that automatically identify the contextual polarity for a large set of sentence expressions, achieving results that are significantly better than baseline.

Vu Dung Nguyen, Blesson Varghese and Adam Barker [20] made an analysis of information retrieved from micro blogging services such as Twitter can provide valuable insight into public sentiment in a geographic region. This insight can be enriched by visualizing information in its geographic context. Two underlying approaches for sentiment analysis are dictionary based and machine learning. The former is popular for public sentiment analysis, and the latter has found limited use for aggregating public sentiment from Twitter data. The research presented in their paper aimed to widen the machine learning approach for aggregating public emotion. A framework for analysing and visualizing public sentiment from a Twitter corpus was developed. In the same paper, a dictionary-based approach and a machine learning approach were implemented within the framework and compared using one UK based case study – the royal birth of 2013.

Tumasjan, Sprenger, Sandner and Welpe [21] forecasted election results using Twitter data, and showed how it can be predicted using the tweets by various people. Saif, He, Alani also have done sentiment analysis over Twitter organizations a fast and effective way to monitor the public's feelings towards their brand, business, directors etc. [22]

Twitter analysis has always been an important tool to build predictive models. Using tweets and micro-blogging posts Wang, Can, Kazemzadeh, Bar, Narayanan described public mood towards 2012 US Election. [23]

Nasukawa and Yi [24] identify local sentiment as being more reliable than global document sentiment, since human evaluators often fail to agree on the global sentiment of a document. They focus on identifying the orientation of sentiment expressions and determining the target of these sentiments. Shallow parsing identified the target and the sentiment expressions; the latter was evaluated and associated with the target. Their system also analysed local sentiments but aims to be quicker and cruder. They charge sentiment to all entities juxtaposed in the same sentence as instead of a specific target.

In "Sentiment analyzer: Extracting sentiments about a given topic using natural language processing techniques" [25], the authors follow up by employing a feature-term extractor. For a given item, the feature extractor identified parts or attributes of that item. e.g., battery and lens are features of a camera.

B. Efficient Market Hypothesis

In the literature there is an existence of efficient market hypothesis (EMH) which states at any point of time and in a liquid market security prices fully reflect all available information. It further categories market into three degrees: weak, semi-strong and strong, which addresses the inclusion of non-public information in market prices. [26] [27] [28] [29]

This theory contends that since markets are efficient and current prices reflect all information, attempts to outperform the market are essentially a game of chance rather than one of skill. In the weak form of EMH, future prices cannot be predicted by analysing prices from the past. The semi-strong form of EMH assumes that current stock prices adjust rapidly to the release of all new public information. It contends that security prices have factored in available market and non-market public information. The strong form of EMH assumes that current stock prices fully reflect all public and private information. It contends that market, non-market and inside information is all factored into security prices and that no one has monopolistic access to relevant information. It assumes a perfect market and concludes that excess returns are impossible to achieve consistently.

The research reported in this paper is motivated towards analysing public sentiment related to the stock of different companies in real-time and rapidly correlating it with the original stock price changes for a period of 4 weeks. In this paper we studied and analysed the news sources available in public domain that influences the market and the public coconsciousness for a period of 4 weeks to study how the market moves on day to day basis with the available news using dictionary based approach. Public sentiment towards a particular company, for example, positive, negative or neutral, is understood by comparing news against lexicons from dictionaries. And these moods finally build the decision of selling, buying or holding the particular stock in order to have maximized profit out of every stock of the particular company. The positive mood pertaining to a company motivates people to invest more in a particular company over the others with the hope that the company where the person is going to invest is performing well, and the person will get profit in return. On the other hand, the negative and neutral mood channelizes the money to a company which has more positivity in the market place. Hence, the resultant positive mood finally becomes the index which determines the movements of stock prices of the particular company.

III. METHODOLOGY

This paper proposes a real time prediction model of sentiment analysis and news analytics. And also, over a period of 4 weeks, we compare and analyse our predicted sentiments with the market trends of 15 different companies namely – Wipro (WIPRO), Infosys (INFY), Unilever NV (UN), Tata Consultancy Services (TCS), International Business Machines (IBM), Tech Mahindra (TECHM), Reliance Industries Limited (RIL), Nestle India (NESTLEIND), General Motors (GM), Bharat Heavy Electricals Limited (BHEL), Google Inc (GOOG), Hindustan Motors (HINDMOTORS), Apple Inc (AAPL), Amazon (AMZN), IDBI Bank (IDBI) to find correlation between predicted sentiment curve with stock price trends and to test the Efficient Market Hypothesis in the market.

The prediction algorithm of this model to predict stock trends is driven by 3 indicators: Positive, Neutral and Negative.

The index of degree of positivity or the degree of negativity or the degree of neutrality is scored from the result of analysis of the available relevant real time news filtered and classified from a large stack of news collated from various popular news sources and press releases by the particular organizations.

This predictive news mining methodology comprises of 6 major components, namely the Database, the News Sources, the Collector, the Keywords, the Classifier, the Estimator, the Analyser and the Visualiser.

A. Database

The project required 6 databases to keep different set of data stored. T_1 is the database of all the real time news sources. The headlines and press releases are collected in database T_2. T_3 is the database of classified relevant real time news headings which is the output of the Classifier. T_4 is the output of estimated results which are degree of positivity, degree of negativity and degree of neutrality. T_5 is the storage of analysed and estimated effective degree of positive sentiment of the time span of 4 weeks. And T_6 keeps the keyword.

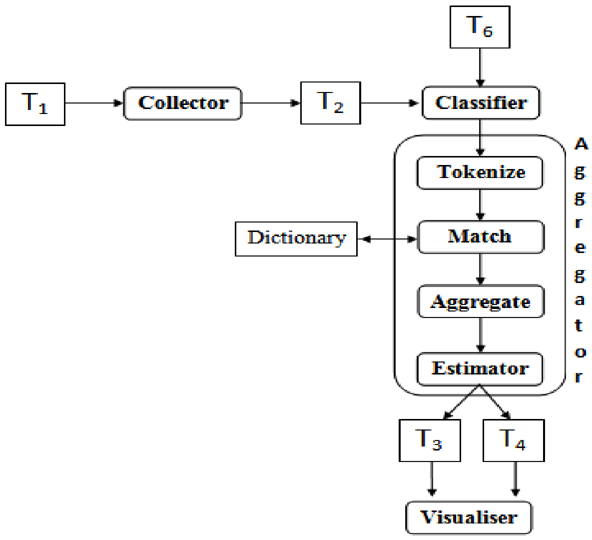

Fig. 1. Flowchart of the methodology components

B. News Sources

this study the data have been collected from the press releases by the respective companies and the regular financial news sources – Wall Street Journal, Financial Times, Forbes, Reuters, CNBC, NDTV, Economic Times, Hindustan Times, Times of India, India Times, Telegraph, CNN Money, Market Watch, and Fortune.

C. Collector

The collector is responsible for collecting the corpus. In this study, the corpuses are the real time news headlines and press releases. The collector extracts the news headlines of the news RSS feeds which are stored into a database (T_1) which nothing but an XML file and finally stores the collected corpus inside another database (T_2). Everyday an average of around 2650 real time business news have been collected to do the sentiment analysis of the companies.

D. Keywords

The keywords are nothing but the words or triggers depending on which the news headlines or the press releases get classified. In our case the keywords will be the name, symbol of the company, the issues and phrases that concerns the company. The filtered news is the news relevant to the keywords. Keyword stays stored in a database (T_6).

E. Classifier

The responsibility of the classifier is to trim the corpus offline. Parser trims the real time news headlines and press releases which are relevant to the particular organization and stores them to another database (T_3) from the collected news stored inside database (T_2). The classifier trims the news headlines on the basis of the. Output of the parsed files is used to analyse and estimate the indicators.

F. Analyser

The principal responsibility of the Analyser is to perform sentiment analysis on the parsed sentences which is the data stored in the T2 database. The approach that has been followed in this paper is dictionary-based approach to analyse sentiment of the relevant news headlines in order to get the indicators. [30] [31] Dictionary based approach or statistic based approach has a significant high percentage of accuracy level and low percentage of tie level. [32]

The analyser which brings out sentence-scores follows an algorithm which calculates cumulative sentiment scores of words present in each line of the parsed sentences. The algorithm which runs for every sentence to calculate its score is comprised of 3 steps – tokenizing, matching, and aggregating.

1) Tokenizing: The news headlines are tokenized using a lexical analyser using R. In T_2 database, a vector of sentences has been gotten as a result of parsing, from where we had to get an array of scores which gets estimated in the later phase.

The algorithm to tokenize each sentence is as followed:

a) Cleaning up the sentence: The punctuations, control characters, and digits of the parsed sentence were cleaned up and stemmed.
b) Change sentence to lower case: Sentence then is converted into lower case.
c) Split the elements into character vector: The cleaned up lower cased sentence is split into words using.
d) Flatten lists: To produce a vector which contains all the atomic components which occur in the list, it gets flattened or unlisted.

2) Matching: Since here in this paper, we aim at analysing it using the most used dictionary based approach; the next step is to match the words of each sentence with the dictionary. The dictionary has two sub dictionaries of positive and negative words.

a) Match the words of the sentence with the dictionary: For each sentence, each word gets matched with the dictionary.
b) Return TRUE or FALSE: If there's a match between the word of the sentence and the word of the positive or negative dictionary it returns TRUE, otherwise FALSE.
c) Sentence Scoring: The total sentiment score is calculated using the following formula and the sentence score is stored into an array.

sentiment_score = ∑ positive_matches - ∑ negative_matches

3) Aggregating sentence scores: Once each sentence gets a score, the number of sentences with same score gets counted and put along with the particular sentence score. It gets stored into a data frame with Number of sentences with a unique score and the unique score.

Number of sentences with a unique score = ∑ sentences with the unique score.

G. Estimator:

The estimator estimates the sentence wise scores and brings out the degree of positivity, negativity and neutrality in percentage. The estimator follows the following algorithm to estimate the indicators. The formulas that the estimator follows are:

*Total Positive Indicator = ∑ positive score*number of sentences with a particular positive score*

*Total Negative Indicator = ∑ (-1)*negative score*number of sentences with a particular negative score*

Total Neutral Indicator = number of sentences with a sentence score 0.

Total Score = Total Positive Indicator + Total Negative Indicator + Total Neutral Indicator

*Degree of Positivity = (Total Positive Indicator / Total Score)*100 %*

*Degree of Negativity = (Total Negative Indicator / Total Score)*100 %*

*Degree of Neutrality = (Total Neutrality / Total Score)*100%*

H. Visualiser

The result of predictive analysis of a particular time stamp is shown in a pie chart. The pie chart comprises of 3 segments: the degree of positivity, the degree of negativity and the degree of neutrality.

The predicted sentiments of effective positive index have been plotted for 15 companies over a period of 4 weeks. The graph of original stock market movement and the movement of predicted sentiment result have been compared in order to find out co-relation between the original curve and predicted curve and also to test efficient market hypothesis.

1) Wipro(13-03-2014 to 5-04-2014):

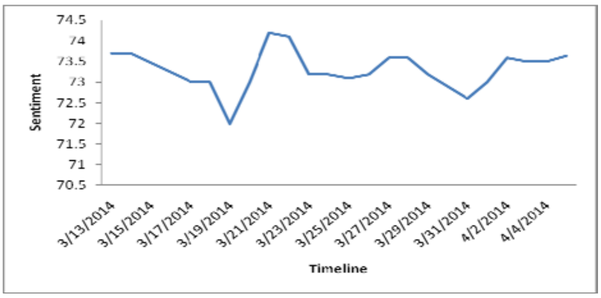

Fig. 2. Curve generated from real time news sentiment of Wipro.

Fig. 3. Original stock price curve of Wipro

2) Infosys(25-03-2014 to 25-04-2014):

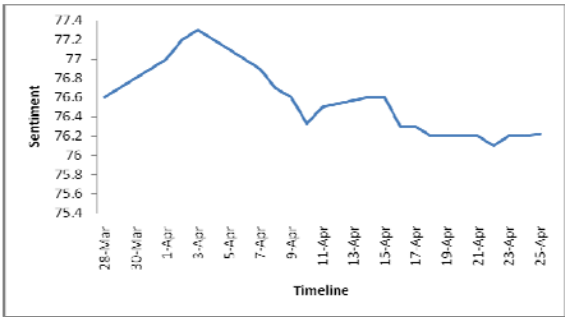

Fig. 4. Curve generated from real time news sentiment of Infosys

Fig. 5. Original Stock Price Curve of Infosys

3) Unilever(31-03-2013 to 25-04-2014):

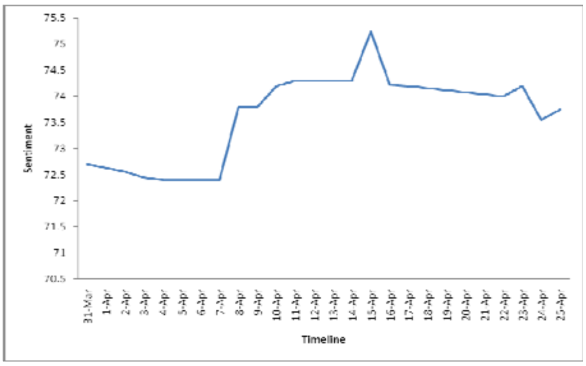

Fig. 6. Curve generated from real time news sentiment of Unilever

Fig. 7. Original Stock Price Curve of Unilever

4) Tata Consultancy Services (25-3-2014 to 24-4-2014):

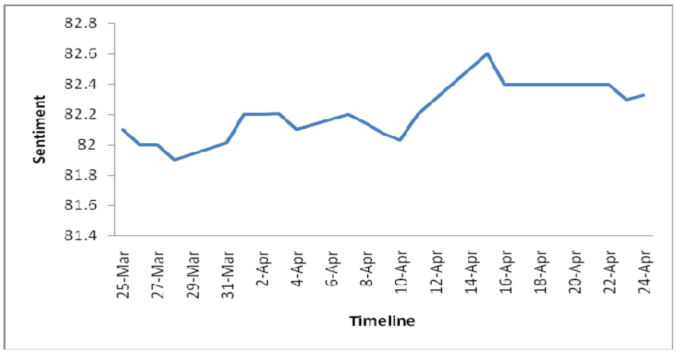

Fig. 8. Curve generated from real time news sentiment of TCS

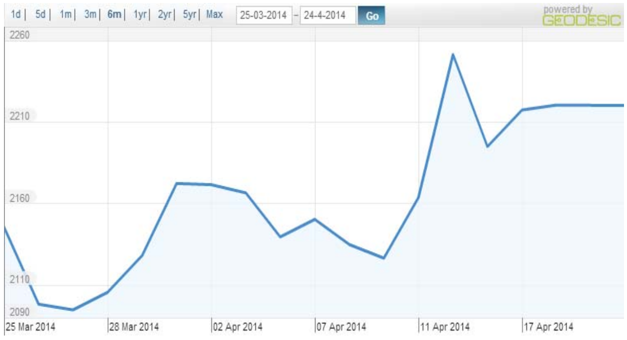

Fig. 9. Original Stock Price Curve of TCS

5) International Business Machines(28-3-2014 to 25-4-2014):

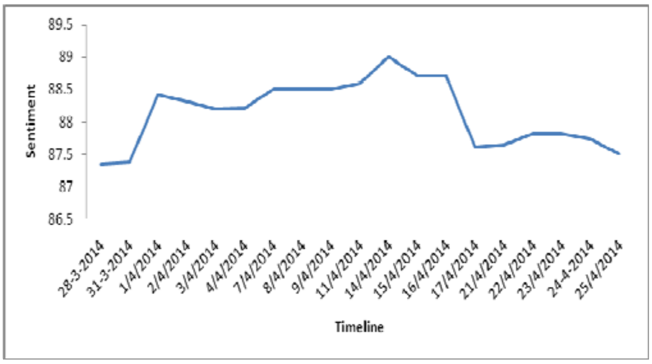

Fig. 10. Curve generated from real time news sentiment of IBM

Fig. 11. Original stock price curve of IBM

6) Tech Mahindra

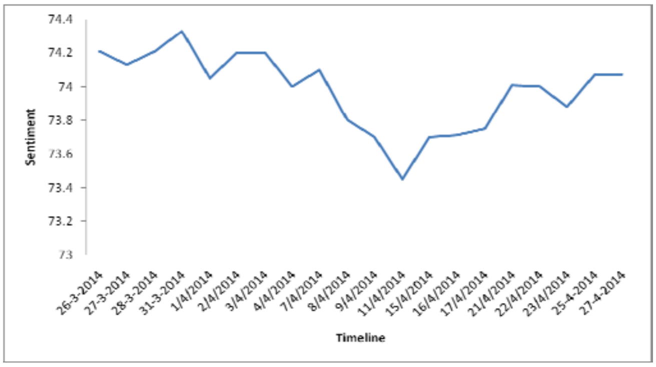

Fig. 12. Curve generated from real time news sentiment of Tech Mahindra

Fig. 13. Original stock price curve of Tech Mahindra

7) Reliance Industries (25-03-2014 to 25 -04-2014):

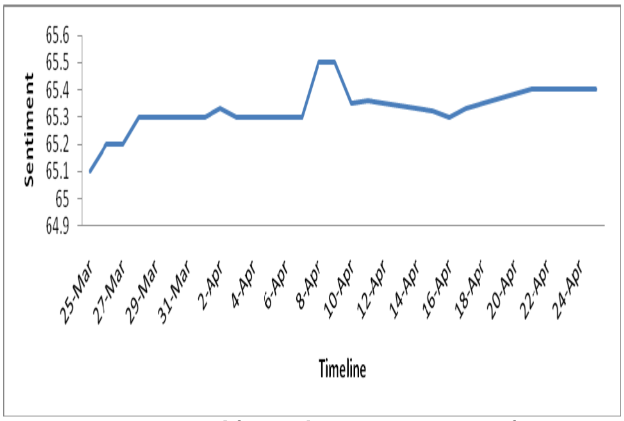

Fig. 14. Curve generated from real time news sentiment of RIL.

Fig. 15. Original curve of Reliance Industries Ltd.

8) Nestle India:

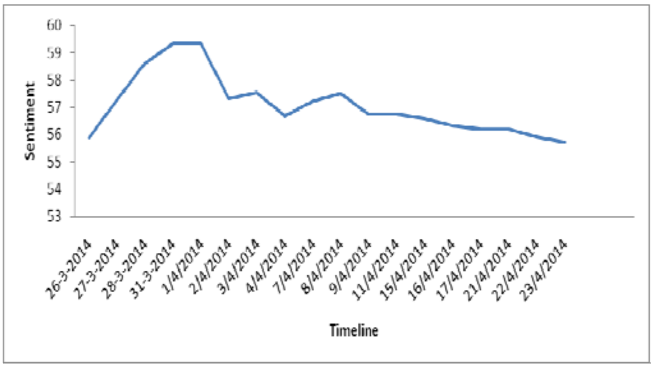

Fig. 16. Curve generated from real time news sentiment of Nestle India

Fig. 17. Original curve of Nestle India.

9) General Motors Company:

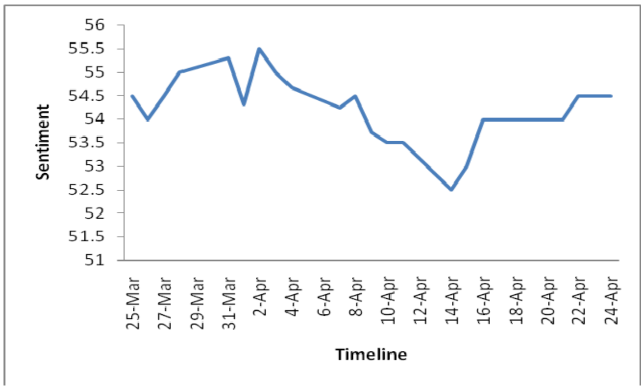

Fig. 18. Curve generated from real time news sentiment of GM Company.

Fig. 19. Original curve of GM Company.

10) Bharat Heavy Electronics Ltd. (BHEL):

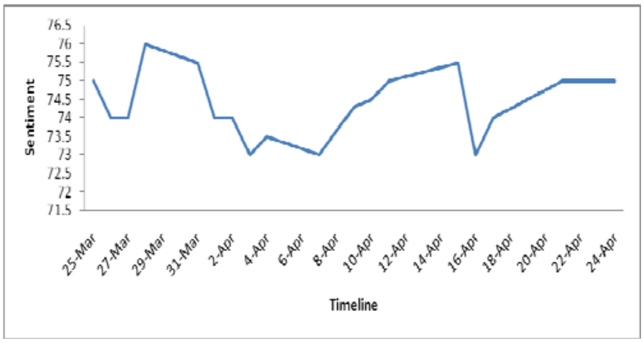

Fig. 20. Curve generated from real time news sentiment of BHEL

Fig. 21. Original stock price curve of BHEL

11) Google Inc:

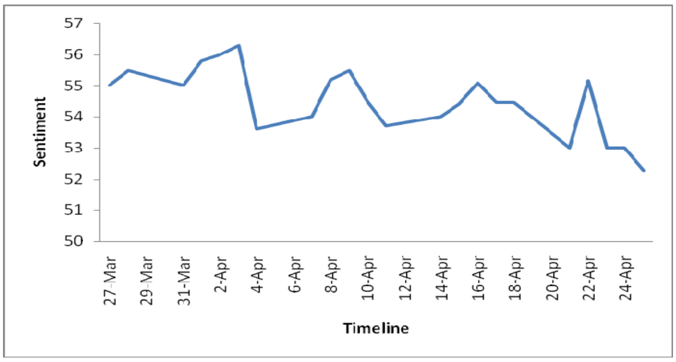

Fig. 22. Curve generated from real time news sentiment of Google Inc.

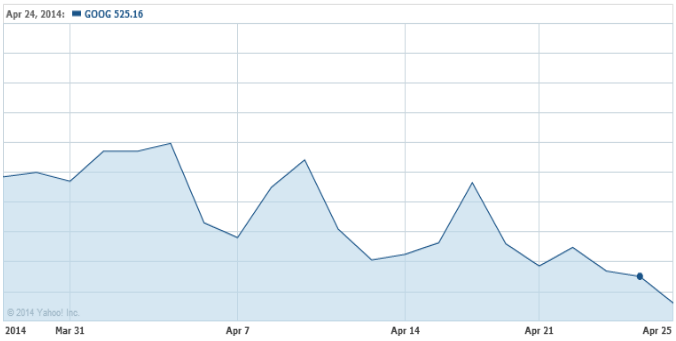

Fig. 23. Original stock price curve of Google Inc.

12) Hindustan Motors

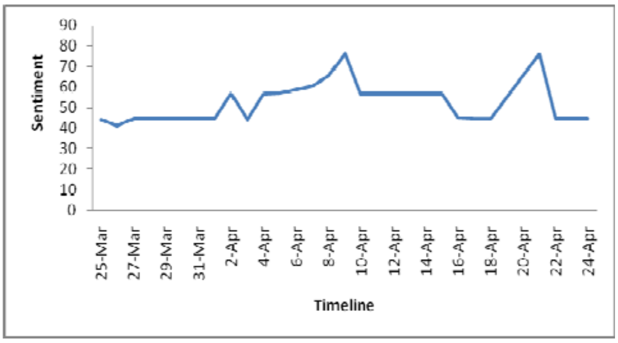

Fig. 24. Curve generated from real time news sentiment of Hindustan Motors

Fig. 25. Original stock price curve of Hindustan Motors

13) Apple Inc:

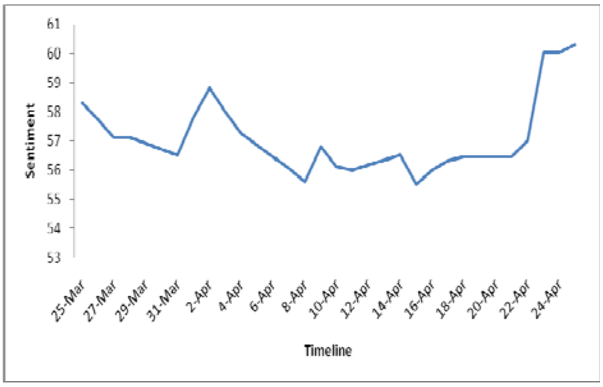

Fig. 26. Curve generated from real time news sentiment of Apple

Fig. 27. Original stock price curve of Apple

14) Amazon

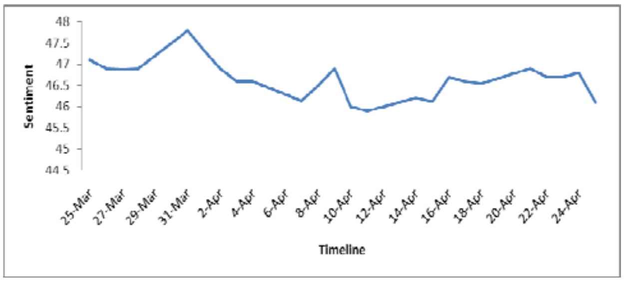

Fig. 28. Curve generated from real time news sentiment of Amazon

Fig. 29. Original stock price curve of Amazon

15) IDBI Bank

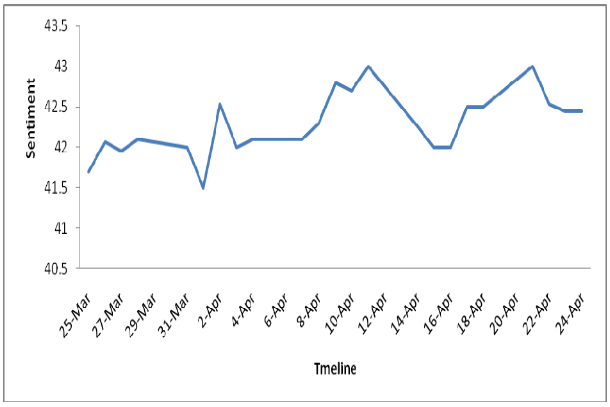

Fig. 30. Curve generated from real time news sentiment of IDBI Bank

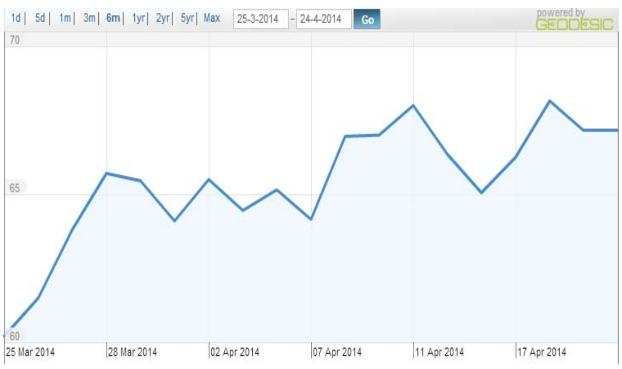

Fig. 31. Original stock price curve of IDBI Bank

IV. EVALUATION

Evaluation of the result was done in order to measure accuracy of the prediction model and also to compare and find out the co-relation between predicted sentiment trends and original stock price movement, if at all any. [33]

In case of all the companies, the news got filtered without leaving out relevant news. The relevancy check of the filtered news from the collected set of news is 100%. The predictive model's resultswere amalysed in terms of Precision, Recall, Specificity, False Discovery Rate, Accuracy, F_1 harmonic factor, Mathew's Correlation Coefficient, Error in prediction, Mean Absolute Error, MAPE, Root Mean Square Error (RMSE), Coefficient of Variation of RMSE.

The outcomes have been formulated in a 2X2 contingency table or confusion matrix with the following components: True Positive (*tp*), False Positive (*fp*), True Negative (*tn*), and False Negative (*fn*) of the real time news headlines of different set of companies classified from news collected from the news sources.

Confusion Matrix		**Target**			
		Positive	**Negative**		
Model	**Positive**	330	144	***Positive Predictive Value***	69.62
	Negative	53	132	***Negative Predictive Value***	71.35
		Sensitivity	***Specificity***	**Accuracy** = 70.1	
		86.16	47.83		

TABLE I 2X2 CONTINGENCY MATRIX

As we test this model on different set of real time news headlines, the error has always been in between 23.5% to 41% and the mean absolute error has been measured around 30%.

The MAPE has been measured as 40.7%. RMSE was measured as 30.3%. Finally, the coefficient of variance of RMSE is 1.009.

Mathew's Correlation Coefficient = 0.4

F_1 harmonic factor of this model is 78.1%

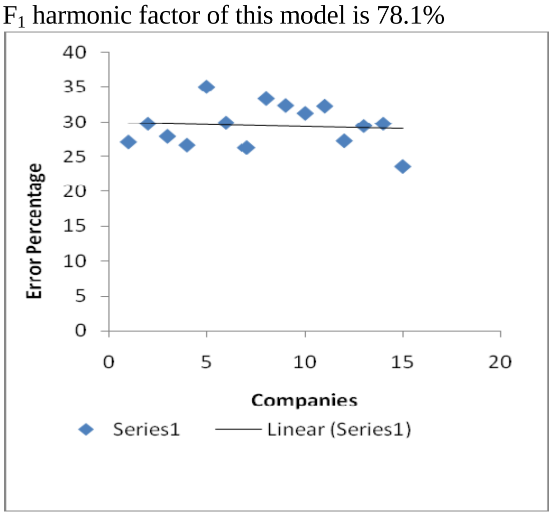

Fig. 32. Error percentages in predicting correct sentiment of different companies, X axis is company's index (in the order they have been mentioned)

The graphs we generate from the predicted tends of sentiment are compared with the movements of original stock curve. And this comparison method yields out a mean correlation of 67.09%. This further passes the effective market hypothesis that says in any liquid market, at any point of time, stock prices reflect all information available in the public domain.

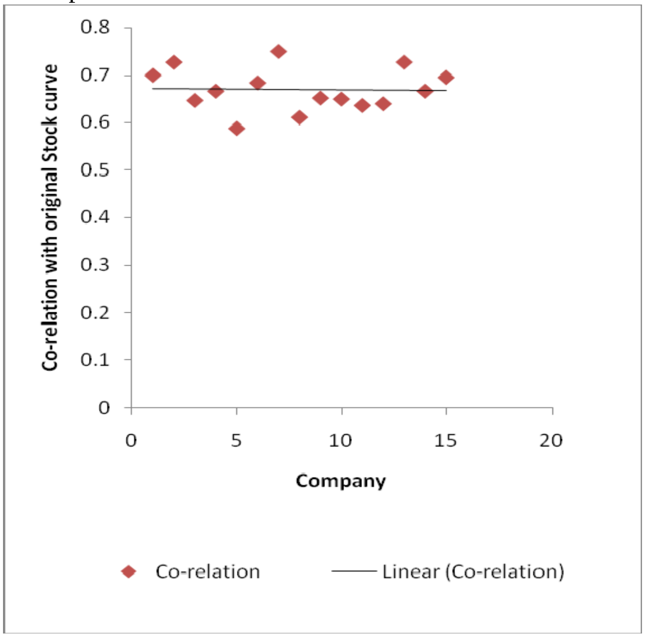

Fig. 33. Correlation between sentiment curve and original curve, X axis is company's index (in the order they have been mentioned)

V. Conclusion And Future Work

As each of the news carries information, news pertaining to a specific company builds perception regarding the organization's policies, growth and performance. All news together builds an overall sense about a particular company. This sense becomes an important fact while trading. A trader invests more when he/she feels that the company's stock price will go up and will gain profit. And more positive the positive index becomes more trust grows and that influences a trader or investors choice of investment and participation. Less positive sentiment will reduce the level of trust and hence the stock price curve will decline.

Every time one runs the prediction model it, shows the current sentiment of the company calculated from relevant news headlines and press releases available at that point of time with an accuracy of 70.1%. The collected set of news is real time news; hence all available news gets collected by this model avoiding all sorts of bias or prejudice. That makes the study more perfect.

The study using the predictive model has collected news at the time of the market closure and then analyzed the sentiment on each company. This tracked sentiment of 4 weeks reveals a strong correlation between the original stock price curve and the curve of effective positive index generated from predictive news mining model that has been presented in this paper. This study reveals that there is a very strong correlation of about 67% between news sentiment and original stock price curve. This easily satisfies the test of effective market hypothesis and shows that the sentiment positively reflects on the stock price movement. A market with such a high index indicates towards existence of a semi-strong to a strong efficient market hypothesis.

Although the accuracy level of this predictive model is satisfactory, it can further be developed by incorporating more complex classifiers, and analysis techniques of machine learning or data mining. The future scope of this paper can be varied. We can compare the sentiment of several competitive companies and show how the sentiment around one company impacts on the other companies over the time. Also we can also study other factors ranging from social to legal issues to check whether at all they have any impact on overall market scenario. Using this proposed model, we can do several financial modelling like portfolio management, risk estimation models, and several other strategic modelling etc. where perception plays an important role.

Acknowledgment

We would like to thank Dr. Debika Bhattacharjee, HOD of Dept of CSE of IEM, Kolkata and other faculties for their support during the study of this paper.

References

[1] Tetlock 2007, "Giving Content to Investor Sentiment: The Role of Media in the Stock Market", Journal of Finance, 62(3).
[2] Da, Engleberg and Gao 2009, In Search of Attention
[3] Odean and Barber 2008
[4] diBartolomeo and Warrick 2005
[5] Mitra, Mitra and diBartolomeo 2009, „Equity portfolio risk estimation using market information and sentiment", Quantitative Finance, 9(8), pp. 887-895;
[6] Dzielinski, Rieger and Talpsepp 2010, Volatility, asymmetry, news and private investors", The handbook of news analytics in finance, Wiley Finance
[7] Antweiler, W. & Frank, M.Z. 2004, „Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards", The Journal of Finance, 59(3), pp. 1259-1294
[8] Atje, R. & Jovanovic, B. 1993, "Stock markets and development", European, Economic Review, 37(2-3), pp. 632-640
[9] Barber and Odean ‚All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors: The Handbook of News Analytics in Finance Chapter 7
[10] Bamber, Barron, and Stober 1997; Karpoff, 1987; Busse and Green, 2004
[11] Leela Mitra, gautam Mitra, Applications of news analytics in finance:A review : The Handbook of News Analytics in Finance , Willy Finance, 2011, Chapter 1
[12] Mittermayer, M. and Knolmayer, G. 2006, Text mining systems for market response to news: A survey.
[13] Pang, B., Lee, L., Vaithyanathan, S. : Sentiment classification using machine learning techniques. In: Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP). (2002) 79-86
[14] Pang, B., Lee, L.: A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In: Proceedings of the ACL. (2004) 271-278
[15] Singh, V.K., Piryani R., Uddin A., Waila, P, "Sentiment analysis of movie reviews: A new feature-based heuristic for aspect-level sentiment classification," published in Automation, Computing, Communication, Control and Compressed Sensing (iMac4s), 2013 International Multi-Conference on 22-23 March 2013, pp. 712-717
[16] V. K. Singh, P. Waila, Marisha, R. Piryani & A. Uddin, "Sentiment Analysis of Textual Reviews: Evaluating Machine Learning, Unsupervised and SentiWordNet Approaches", In Proceedings of 5th International Conference of Knowledge and Smart Technologies, Burapha University, Thailand, Jan. 2013
[17] V. K. Singh, R. Piryani, A. Uddin & P. Waila, "Sentiment Analysis of Movie Reviews and Blog Posts: Evaluating SentiWordNet with different Linguistic Features and Scoring Schemes", In Proceedings

of 2013 IEEE International Advanced Computing Conference, Ghaziabad, India, Feb. 2013.
[18] T. Willson, J. Wiebie, P. Hoffman, “Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis”
[19] Grefenstette, Y. Qu, J.G. Shanahan, and D.A. Evans. 2001.Coupling niche browsers and affect analysis for an opinion mining application. In RIAO-2004.
[20] Vu Dung Nguyen, Blesson Varghese, Adam Barker, “The Royal Birth of 2013: Analysing and Visualising Public Sentiment in the UK Using Twitter”
[21] A. Tumasjan, T. O. Sprenger, P. G. Sandner and I. M. Welpe, “Election Forecasts With Twitter: How 140 Characters Reflect the Political Landscape,” Social Science Computer Review, Vol. 29, No. 4, 2011, pp. 402-418.
[22] Saif H, He Y, Alani H, "Semantic Sentiment Analysis of Twitter", 11th International Semantic Web Conference, Boston, MA, USA, November 11-15, 2012, Proceedings, Part I, pp. 508-524
[23] Wang H, Can D, Kazemzadeh A, Bar F, Narayanan S, "A system for real-time Twitter sentiment analysis of 2012 U.S. presidential election cycle", In proceeding of ACL '12 Proceedings of the ACL 2012 System Demonstrations, pp. 115-120
[24] Nasukawa, T., Yi, J.: Sentiment analysis: Capturing favorability using natural language processing. In: The Second International Conferences on Knowledge Capture. (2003) 70-77
[25] J. Yi, T. Nasukawa, R.B., Niblack, W.: Sentiment analyzer: Extracting sentiments about a given topic using natural language processing techniques. In: 3rd IEEE Conf. on Data Mining (ICDM'03). (2003) 423-434
[26] Michael C. Jensen, Some Anomalous Evidence Regarding Market Efficiency, 1978 Journal of Financial Economics, Vol. 6, Nos. 2/3 (1978) 95-101
[27] Burton G. Malkiel, Reflections on the Efficient Market Hypothesis: 30 Years Later, 2004 Meetings of the Eastern Finance Association in Mystic, Connecticut
[28] Malkiel, B.G., 2003. The efficient market hypothesis and its critics, Journal of Economic Perspectives 17, 59–82
[29] Basu,S., Investment Performance of Common Stocks in Relation to Their Price-Earnings Ratios: A Test of the Efficient Market Hypothesis
[30] Minqing Hu and Bing Liu. "Mining and Summarizing Customer Reviews." Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2004), Aug 22-25, 2004, Seattle, Washington, USA.
[31] Bing Liu, Minqing Hu and Junsheng Cheng. "Opinion Observer: Analyzing and Comparing Opinions on the Web." Proceedings of the 14th International World Wide Web conference (WWW-2005), May 10-14, 2005.
[32] B. Pang, L. Lee, S. Vaithyanathan, “Thumbs up? Sentiment Classi¯cation using Machine Learning Techniques,” in the proceedings of the conference of Empirical Methods in Natural Language Processing (EMNLP), pp 79 – 86, 2002
[33] The original stock price curves have been taken from Money Control (http://www.moneycontrol.com/) and Yahoo Finance (https://in.finance.yahoo.com/).