Prescribing the Stock Market using Web Media

J. Lavanya[1] R.K.Shiva Ranjani[2] N. Sree Monisha[3] MS. G. Indumathi[4]

[1,2,3]U.G Student [4]Assistant Professor

[1,2,3,4]Department of Computer Science & Engineering

[1,2,3,4]Velammal Institute of Technology, Panchetti, Tamil Nadu, India

Abstract— Analyze the financial data for making a prediction in stock markets by using big data analytics and recommend the user investment in various category stock markets based on web media. Hadoop is a Java-based programming framework that supports the processing of large data sets in a Parallel and distributed computing environment. It makes use of the commodity hardware Hadoop is Highly Scalable and Fault Tolerant. Hadoop runs in the cluster and eliminates the use of a Supercomputer. Hadoop is used as a big data processing engine with a simple master- slave setup. In the stock market industry, predictive analytics plays a key role in providing data-driven decisions for managing the resources under an NSE and BSE. There are certain attributes that trigger redemption by investors are complex in nature to identify and analyze. Predicting redemption behavior requires a sophisticated platform that can capture multiple factors that affect redemption behavior. However, big data predictive analytics using advanced analytics platform can analyze these massive amounts of transaction data and other time trend variables at a macro level. This platform can investigate these factors for near real- time data and can provide highly accurate predictions for the redeeming investors in the future at an investor-level.

Key words: Stock Market, Analysis, Hadoop, Web Media

I. INTRODUCTION

A stock market is the aggregation of buyers and seller of stocks (also called shares), which represent ownership claims on businesses. These could embody securities listed on a public stock exchange in stock market, as well as stock that is only traded privately. Examples of the shares of private companies which are sold to investors through equity crowd funding platforms. Stock exchanges list shares of common equity as well as various security types, e.g. corporate bonds and a convertible bond. The workings of the stock market exchange can be a great confusion for several people. Some people believe in investing in a stock market is a form of gambling and feel that if you invest, you will more than likely end up losing your money.

India has a population of around 1.3 billion, out of which only 20-27 million people invest in shares or mutual funds. We have been touched solely 2% of India and we have to reach the remaining 98%. Since markets are volatile, people are afraid of investing in stocks. Even when they want to invest in it, a lot of mis-selling happens. People with a very low understanding of derivatives are trading in them. We have a market that's not safe for investors to an oversized extent. Newcomers and gullible individuals have issues in understanding the products and end up choosing high-risk products. Once they lose money, they may leave the market forever. This is the main reason for the low percentage of people investing in the stock market in India comparing to other countries.

Existing research has shown the utility of public sentiment in social media across a wide range of applications. Several works showed social media as a promising tool for stock market prediction. There are many social media platforms such as Twitter, Blogs, Discussion Board, News and there are many other forms of social media are present to share their thoughts, information and their own opinion about the particular news. This has played a major role in determining the people mind swing in predicting the stock market. There is much analysis in predicting the stock market How to invest, Where to invest, when to invest. These are the important question that arises in people mind when they think of investing in the stock market. Also, there are a lot of papers explaining the relationship between social media and the stock market. In this paper, it explains the way to find the accuracy in predicting the stock market is analyzed. Our experimental results really show how our analysis meant improve the prediction of the stock market.

II. LITERATURE SURVEY

Lily Fang and Joel Peress[1] proposed that Media can affect directly or indirectly the price of stocks in stock market exchange. How it because of the investigation of this hypothesis by studying the cross-sectional relation between media coverage and expected stock returns. It finds that stocks with no media coverage earn higher returns than stocks with high media coverage even after controlling for well-known risk factors. These results are more pronounced among small stocks and stocks with high individual possession, low analyst following, and high idiosyncratic volatility. In this, findings suggest that the breadth of information dissemination affects stock returns. In this paper, it does not process on the effectiveness of the sentiment analysis in the stock prediction task via a large scale experiment.

Thien Hai Nguyena, Kiyoaki Shiraia, Julien Velcinb[2] presents Stock price forecasting is important in predicting the movements of stock markets. There is some research trying to extract sentiments from feedbacks such as reviews and comments from people through the internet. It mainly extracting mood information through sentimental analysis on social media data. Then, these sentiments will be integrated to build a model to predict the stock exchange. One contribution feature of this was topic-sentiment to improve the prediction of the stock market. The support vector machine has long been recognized as been able to efficiently handle high dimensional data and has been shown to perform well on classification. It uses with linear Kernel as the prediction model. This study does not explain the predictive relationships between social media and firm equity value and the relative effects of social media metrics.

Xueming Luo, Jie Zhang, Wenjing Duan[3] proposed that the efficient market hypothesis states that new information may change market expectations and thus move

a firm's stock prices. Financial studies also suggest the notion of information asymmetry in the stock market. To overcome this asymmetry, investors seek additional sources of information beyond sales to determine firm equity worth. The results derived from vector autoregressive models that recommend media-based metrics (Web blogs and client ratings) are significant leading indicators of firm equity worth. Interestingly, typical online behavioral metrics (Google searches and Web traffic) are found to have a significant however substantially weaker forecasting the relationship with firm equity value than social media metrics. A time-series technique, namely, VARX. This modeling approach permits us to capture dynamic interactions and feedback effects. These findings are robust to a standardized set of volume-based measures. It will not determine the predictability of financial markets and cause huge gains or losses.

Arman Khadjah Nassirtoussi, Saeed Aghabozorrgi, Teh Ying Wah, David Chec Ling[4] presents text mining, additionally stated to as text data mining, roughly similar to text analytics, is the method of deriving high-quality data from text. High-quality info is usually derived through the devising of patterns and trends through means such as statistical pattern learning. In this, comparative analysis of each system expands into the theoretical and technical foundation. This work helps to structure the emerging field and identify the exact aspects which require further research. This does not review the related works that are about market prediction based on online text- mining and produce a picture of the generic components that they all have.

Wesley S. Chan[5] examined the returns to a subset of stocks after public news about them is released. I compare them to different stocks with similar monthly returns, but no identifiable public news. There is a major difference between come back patterns for the two sets. The patterns are seen even after excluding earnings announcements, controlling for potential risk exposure, and other adjustments. It also finds some evidence of reversal after extreme price movements that are unaccompanied by the public. There are no speech tags - so as to identify the most effective method that will allow for the classification of racist documents on the web.

Robert P. Schumaker and Hsinchun Chen[6] says Stock Market prediction has always had a certain appeal for researchers. While various scientific research makes an attempts to find a method but no method has been discovered to accurately predict stock price movement. The difficulty of prediction lies within the complexities of modeling market dynamics. This project learning approach for financial news articles analysis using several different textual representations. This will not solve this problem by comparing the behaviors of investors with access to different media coverage of the same information event.

Joseph Engeleberg and Christopher A. Parson's[7] Disentangling the causative impact of media news from the impact of the events being reported is challenging. A number of recent studies demonstrate sturdy correlations between stories reported by the media and stock market reactions. This paper addresses the causal relationship between the two. Specifically, the question is whether the media coverage of a financial event can alter investor behavior. This paper does not the practice of Retweeting as a way by which participants can be "in a conversation".

III. PROPOSED SYSTEM

The exploration of media-mindful stock developments started with money-related reports and news articles. With the prevalence of Web 2.0, new media sources, for example, web journals, tweets/miniaturized scale websites, exchange sheets, and social news, have developed and assumed imperative jobs in influencing financial exchanges. As contemplate, it found that the feelings of tweets influenced stock patterns for a short period after the arrival of the tweets. As opposed to conventional news, web-based life enables clients to express their suppositions and sentiments through remarks. Such client commitment productively improves data dispersal and builds the estimation of the data. We are utilizing Google API to associate with the web and get the news articles to investigate the stock patterns and prescribed the right stocks to the clients.
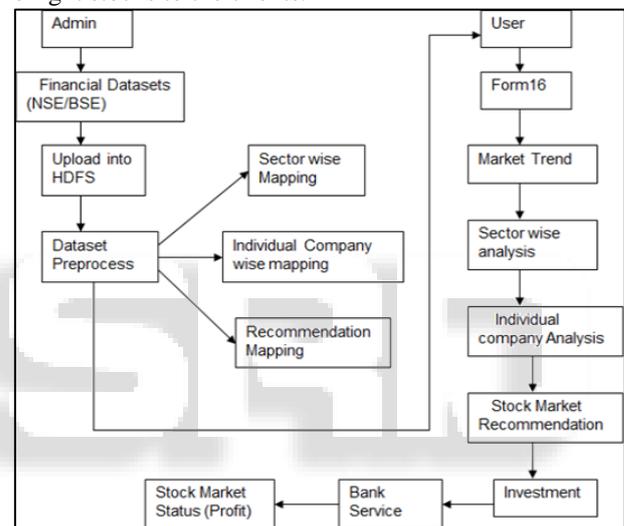
A. Big data environment

Tremendous Collection of information is recovered from open source datasets that are freely accessible from real Application Providers like Money Control. Enormous Data Schemas were investigated and a Working Rule of the Schema is resolved. The CSV (Comma isolated qualities) and TSV (Tab Separated Values) documents are stored in HDFS (Highly Distributed File System) and were perused Master and controlled utilizing Java API that itself created by us which is amicable, light weighted and effectively modifiable.

B. Preprocessing and User Form16

A preprocessing is a backend work running in Hadoop clusters and furthermore called as long- running employment as it is planned to process mass information with the goal that the application would make utilization of the outcomes created for refreshing. Dataset mapping process done in the preprocessing stage the whole part of both NSE and BSE organizations will be mapped date astute, month insightful and year shrewd information. The hazard factor of every segment and friends will be finished over the preprocessing time. The client financial exchange speculation dependent on the client qualification, for example, aadhar number, Account

subtleties, yearly pay, working status,conjugal status, credit status, and so forth every one of these fields will be accessible on form16.

C. Market trends

Our application gives a history of market information in NSE and BSE, with the goal that the client will have an idea regarding market pattern shared assets. We have a different division such as Finance, Management Information Systems, and Computer Science. A client can see the historical backdrop of all the area by utilizing our forecast component among the segments we have diverse organizations client can see the history of every organization in the premise of date, month and year wise.

D. Recommendation Based On Web Media

In light of Web Media substance, for example, News articles and internet-based life and talk board where there will be articles about the stocks which perform well and furthermore about different stocks will be there where we use Google API to separate the data from the web and dependent on that we prescribe the stocks to the clients dependent on hazard factors and taking into all contemplations.

IV. Conclusion

From the ancient age of transmitting information via smoke signals to the telegraph era to the internet era, information, as an important market factor and price factor, constantly influences and reconstructs financial markets. We endeavored to perform a quantitative analysis to understand, in detail, the mechanisms by which information diffuses through the web and its impact on stock markets from the perspectives of investors cognitive behaviors, corporate governance, and stock exchange regulation. With the recent advancements in computational power, it is possible to quantify the influences of Web media on financial markets, allowing us to better understand the mechanisms of markets, which will be able to protect investors with the most valuable information, assist in corporate management, and provide decision makers with the most reliable inputs for the health of the market.

Reference

[1] L. FANG AND J. PERESS, "MEDIA COVERAGE AND THE CROSS-SECTION OF STOCK RETURNS," J. FINANCE, VOL. 64, NO. 5, PP. 2023–2052, 2009.

[2] T. H. Nguyen, K. Shirai, and J. Velcin, "Sentiment analysis on social media for stock movement prediction," Expert Syst. Appl., vol. 42, no. 24, pp. 9603–9611, 2015.

[3] X. Luo, J. Zhang, and W. Duan, "Social media and firm equity value," Inf. Syst. Res., vol. 24, no. 1, pp. 146–163, 2013.

[4] A. K. Nassirtoussi, S. Aghabozorgi, T. Y. Wah, and D. C. L. Ngo, "Text mining for market prediction: A systematic review," Expert Syst. Appl., vol. 41, no. 16, pp. 7653–7670, 2014.

[5] W. S. Chan, "Stock price reaction to news and no- news: Drift and reversal after headlines," J. Finance. Econ., vol. 70, no. 2, pp. 223–260, 2003.

[6] R. P. Schumaker and H. Chen, "Textual analysis of stock market prediction using breaking financial news: The AZF in text system," ACM Trans. Inf. Syst., vol. 27, no. 2, pp. 12:1–12:19,2009.

[7] J. E. Engelberg and C. A. Parsons, "The causal impact of media in financial markets," J. Finance, vol. 66, no. 1, pp. 67–97, 2011.

[8] V. Lavrenko, et al., "Language models for financial news recommendation," in Proc. 9th Int. Conf. Inform. Knowl. Manage. 2000, pp. 389–396.

[9] M.-A. Mittermayer and G. F. Knolmayer, "Newscast: A news categorization and trading system," in Proc. IEEE 6th Int. Conf. Data Mining, 2006, pp. 1002–1007.

[10] T. Preis, H. S. Moat, and H. E. Stanley, "Quantifying trading behavior in financial markets using Google trends," Sci. Rep., vol. 3, pp. 1–6, 2013.

[11] S. Y. Yang, S. Y. K. Mo, A. Liu, and A. A. Kirilenko, "Genetic programming optimization for a sentiment feedback strength based trading strategy," Neurocomputing, vol. 264, pp. 29–41, 2017.

[12] H. S. Moat, et al.,"Quantifying Wikipedia usage patterns before stock market moves," Sci. Rep., vol. 3, pp. 1–5, 2013.

[13] C. Curme, T. Preis, H. E. Stanley, and H. S. Moat, "Quantifying the semantics of search behavior before stock market moves," Proc. Natl. Acad. Sci., vol. 111, no. 32, pp. 11 600–11 605, 2014.

[14] M. Z. Frank and W. Antweiler, "Is all that talk just noise? The information content of internet stock message boards," J. Finance, vol. 59, no. 3, pp. 1259– 1294, 2004.

[15] S.-H. Kim and D. Kim, "Investor sentiment from internet message postings and the predictability of stock returns," J. Econ. Behav. Organ. vol. 107, pp. 708–729, 2014.

[16] A. Urquhart and R. Hudson, "Efficient or adaptive markets? Evidence from major stock markets using very long run historic data," Int. Rev. Financ. Anal., vol. 28, pp. 130–142, 2013.

[17] M. Makrehchi, S. Shah, and W. Liao, "Stock prediction using event-based sentiment analysis," in Proc. IEEE/WIC/ACM Int. Joint Conf. Web Intel. Agent Technol., 2013, pp. 337–342.

[18] A. K. Nassirtoussi, S. Aghabozorgi, T. Y. Wah, and D. C. L. Ngo, "Text mining of news-headlines for FOREX market prediction: A multi-layer dimension reduction algorithm with semantics and sentiment," Expert Syst. Appl., vol. 42, no. 1, pp. 306–324, 2015.

[19] R. P. Schumaker and H. Chen, "A quantitative stock prediction system based on financial news," Inf. Process. Manag., vol. 45, no. 5, pp. 571–583, 2009.