

Role of Social Sentiment Analysis in Stock Trends Forecasting

K.Anuratha, M.Parvathy, S.Sujeetha, J.Ghayathri

Abstract: Social media like Face book, Twitter have attracted attention from various sectors of study in recent years. Most of the people share ideas, opinions on various topics such as Stock Market Prediction, Digital marketing, Movie review, Election Results Prediction and Product reviews etc., Forecasting Financial Market is considered to be one of the significant applications of Sentiment Analysis on Social Data like Face book, Twitter. It is essential to accurately predict the movements in stock trends, as the stock market trends are volatile. In the past few years several researches have been carried out for predicting the future trends of stock market through sentiment analysis on social media comments. This paper gives the survey on the various techniques, tools and methodologies adopted by several researchers on Stock Market Prediction based on sentiment analysis of Social networks.

Keywords: Stock Prediction, Twitter, Sentiment Analysis, Classifiers, Accuracy, Deep Learning

I. INTRODUCTION

It is always interesting for the researchers to find ways to predict what will happen in the future. Social media is a communication platform contains valuable knowledge hidden in it. Information available in the social media resembles real world events and they can be exploited by the researchers to enhance the application capabilities. Stock market prediction can be considered as the one of the important applications of social media. The Stock market is a complex system as it is been influenced by the political, economical and social factors. The prices of stock are very dynamic and impressionable to changes due to the nature of financial domain. Though it is a complex system still stock market is one of the important economic factors.(Al-Augby,2015)[17].The focus of stock market forecasters is to develop a successful approach to predict the stock prices. Prediction of stock market is one of the tough tasks because

globe the use of social networks is popular and huge, as it provides a medium to express, share and publish the opinions of people. The effect of social media in stock market prediction has been studied by several researchers, in recent times. Social networks play important role in the society to share the ideas and thoughts of the users through the internet among the virtual community. The knowledgeextracted from the social networks can be applied to predict movement of stock market to some extent.

The most famous micro blog Twitter allows it users to create tweets, short messages that can be shared with and responded by other users of Twitter. The users are much focused on the message they wish to communicate, as twitter employs a restriction on message size. This feature of Twitter makes the tweets good candidates for the Sentiment Analysis task.

Sentiment analysis falls under Natural Language Processing (NLP), a branch of Machine Learning which deals with How computers process and analyze human linguistics?.

This paper is planned as follows: Section II describes Sentiment Analysis. Section III describes the inference from the related research on stock prediction – Stock Prediction Roadmap. The Comparative Study on the different approaches is summarized in Section IV. Section 5 describes the Conclusion on the work carried out and proposes the scope for future work.

II. SENTIMENT ANALYSIS

Sentiment analysis is the process of determining opinion from people's emotion and feelings. Sentiment classification can be done at phrase level, sentence level and document level. The sentiment analysis uses Natural Language Processing (NLP) to divide the language units in to three categories: Negative, Positive and Neutral [20].

The different opinions of people, shared in the social media play significant role in the process of decision making and recommendations [20]. The analysis on micro blogging websites are done using Sentiment Analysis. The contents of Social Media such as posts, tweets, photos are analyzed by people of different community such as politicians, marketers and analysts etc., Nowadays, stock market investment plays an inevitable role in the finance sector, as high stock market value is considered as the parameter of high economies. The volatile nature of stock market has

Revised Manuscript Received on July 22, 2019.

K.Anuratha, Asst. Professor, Dept. of IT, Sri Sai Ram Institute of Technology, Sivaganga (Tamil Nadu), India.

M.Parvathy, Asst. Professor, Dept. of IT, Sri Sai Ram Institute of Technology, Sivaganga (Tamil Nadu), India.

S.Sujeetha, Asst. Professor, Dept. of IT, Sri Sai Ram Institute of Technology, Sivaganga (Tamil Nadu), India.

J.Ghayathri Professor, Dept. of CSE, K.L.N. College of Information Technology, Sivaganga (Tamil Nadu), India.

equal chances for earning money and losing money as well. There is no single technique to predict the stock movement accurately, so there were a lot of experiments done by researchers to get better results. Due to the universal use of social media websites, they can be considered as important in prediction of stock movements, as investors share their opinions and thoughts in the media.

Using Sentiment Analysis as a contributing factor in stock movement prediction is significant field of studies, recent times and below is the presentation of overview of recent work in the past five to six years using Sentiment Analysis based on social media in stock market prediction.

III. STOCK MOVEMENT PREDICTION ROAD MAP

This section gives a detailed survey on the process of stock prediction using various approaches like social information, moods of investors, machine learning models, deep learning models etc.,

A. Stock Trends Prediction on Social Network Information, Psychological states of users and Collective Sentiment Analysis :

The authors Zheng Chen and Xiaoqing Du (2013) [16], in their work developed a unique Social Network Systems Weibo.com and Renren.com to play the roles of Twitter and Face book, as Chinese people don't have access to Face book and Twitter. They have used a unique Chinese Stock forum Guba that has several topics that focus more on stock exchange. They studied the correlation between the stock and its social characteristics that involves their trading behavior over a period of time. They have used a BP-Neural network with Eleven characteristics as input and produced volume and price as output, and has a hidden layer with 20 neurons. They have used only the social media information related to trading but not the opinion analysis of traders because of their belief that the stock investors don't express their emotions online. They concluded that Guba can give a prediction of 56.28% in less than three months and the rate of stock increased to 1.17%.

The Stock Market movement prediction accuracy was improved using Sentiment Analysis in the Porshnev et al (2013)[15]. In this work the researchers aimed at to improving the prediction accuracy through the analysis of Psychological states of Twitter users. They have used Support Vector Machine and Neural Networks Algorithms for prediction on stock market data(S&P500,DJIA) taken from Yahoo Finance website(<http://finance.yahoo.com>). They could achieve the average prediction accuracy of 65% approximately, means the emotional states of the users don't have effect on improving the accuracy.

The researchers Feifei Xu and Vlado Keelj (2014) [14] proposed a method to predict the stock price change for the following day, through collective sentiment analysis for a period of three months from MARCH 2012 to MAY 2012. They have randomly collected the tweets from Stock Twits for the above mentioned duration. The method was implemented in two stages using Natural Language Processing (NLP) and Machine Learning algorithms. In NLP stage they have used Hand-held data with the assistance of Weka Tool Kit, for sentiment analysis. The tweets were classified into positive, negative and neutral based on their polarity. Next sentiment detection process was carried out using three machine learning classifiers

Naive Bayes, Decision Tree (J48 in Weka) and Support Vector Machine (SMO in Weka). They have concluded that they could get the accuracy rates of 71.84% and 74.3% in predicting the movement in stock for the next day, through collective sentiments.

The emotions of market participants can be considered the driving force of prediction of stock movement based on finance researchers' behaviors. These findings were replicated by Michael and Oliver (2015) [11], considering the mood states in approximately 100 million tweets, published in Germany over a period of three years(2011 – 2013) on Twitter. The work was conducted in three steps and they have used Social Mood Index (the sum of positive and negative mood states) and Weighted Social Mood Index (Sum of positive and negative mood followers) for the analysis. They have concluded that there was an increase up to 36 % in a period of 6 months, considering the transaction costs.

Qian et al (2016) [9] aimed at finding the correlation between twitter users' mood changes (approximately 9 days) and their tweets on company stock trends. They have assessed the frequency of twitter posts based on emotions (happy, sad, anger, fear, disgust, and surprise) and retrieved the NASDAQ market closing price to study the relationship on market closing price with mood changes of users.

B. Improvement in stock trends prediction using Machine Learning Models:

Using a Factorization Machine (FM) along with micro blogging sites can improve the prediction of changes in stock trends (Chen et al, 2014) [13]. They have contributed the research in three dimensions. First, insights on why Factorization machine (FM) is better than Generalized Linear Model (GLM) and Support Vector Machine (SVM) were provided based on the relationship of FM with GLM and SVM. Second, explained the benefits of using social media for stock market prediction than other data source and Third, demonstrated the efficiency of FM in prediction. They have concluded that they have conducted the experiment on Sina Weibo (used data of 361 trading days), a popular Chinese Micro blogging service and achieved the prediction accuracy of 81 % with FM, which is more accurate prediction than other models. They also added that the same results can be expected from other social media like Facebook, Twitter.

Authors Yahya and Bayu (2015) [12] have predicted the stock prices of 13 companies of Indonesian Stock Market by simple Sentiment Analysis. They applied sentiment analysis on tweets collected from Twitter Rest API. After the collection of tweets, there were counting of word weights, tagging parts of speech (POS) and sentiment analysis. To classify the tweets and to calculate the sentiments, they have used Naive Bayes Classifier which gave accuracy of 56.5% and Random Forest that accuracy of 60.3%. The classification was supervised with the use of SVM, Neural network, Naive Bayes, Decision tree and random forest. Linear regression methods were used to develop the prediction model. They

have shown the best prediction with 0.9989 and 0.9983 coefficient of determination, using the prediction model that used the sentiments and stock price.

The prediction accuracy and efficiency can be improved by incorporating the specific topics of the organization in to the model used for prediction, developed by Nguyen and Shirai (2015) [10]. They have proposed topic sentiment (sentiments of the company like product, service and so on) feature for improving the prediction accuracy and can be obtained in two ways, first using the (Latent Topic based model) Joint Sentiment / Topic model(JST),second using Aspect Based Sentiment Model. For getting the states of mood on stocks, message boards of 18 stocks were collected from Yahoo Finance Message Board, between July 2012 and July 2013). To handle data at a large scale Support Vector Machine was used to perform classification. The effectiveness of topic-sentiment feature was assessed by comparing the Aspect Based sentiment method with sentiment classification method which used only sentiment information and Latent Dirichlet Allocation (LDA) based method that used only topic information. It was shown that the topic-sentiment feature is better than using only topic feature by 2.14 % and using only sentiments by 2.54 %. They have concluded that the topic- sentiment feature is better than considering only topics or only sentiments.

C. Stock Price Shocks forecasting using Degree of Social Attention (DSA) Framework, Micro blogs and sentiment transfer learning:

A stock price shocks forecasting framework was developed by Zhang et al (2016) [8] that considered the Degree of Social Attention as the factor for stock forecasting. They focused on Chinese stock market – Shenzhen Stock Exchange (SZSE) and Shanghai Stock Exchange (SSE) and the activities in the Chinese social media network Weibo.com. The authors concluded three findings from their work. First, they claimed that the inclusion of Sentiment – Degree of Social Attention (DSA) improved the performance of prediction. Second, among the use of five classifiers for the work, best performed algorithm is Random forest and worst performed is SVM. Third, negative social attention would lead to better forecasting than that of positive.

The use of sentiment analysis and attention indicators that retrieved from micro blogs in forecasting changes in stock movement was studied in “The impact of micro blogging data for stock market prediction: Using Twitter to predict returns, volatility, trading volume and survey sentiment indices” by Nuno et al (2017) [7]. A robust methodology was proposed to assess the values of sentiments in texts and attention indicators to forecast stock market variables. For assessing the values of sentiments for stock market variables five regression models have been tested in the research. They are Multiple Regression (MR), Support Vector Machine (SVM), Random Forest (RF), Neural Network (NN) and Ensemble Averaging (EA) method. The Kalman Filter was used to filter the indicators from micro blogging data acquired from Twitter and the results shown that the micro blogging has significant role in helping stock forecasting.

The skewed news articles distributed among different stocks, lead to low prediction accuracy in the predictions of stock movement because of the few training

samples. This problem was addressed by Xiaodong et al (2018) [5]. A sentimental transfer learning was proposed to transfer the sentimental information learnt from news rich stocks to new poor stocks. They have developed three transfer principles and incorporated them in to source stocks. There were four stages in the research work. First mapping the financial news articles from news – rich stocks and news-poor stocks in to a sentimental feature space. Second, as per instincts on stock trading three principles are developed to guide the process of transfer. Third, candidate stocks generated from each principle were ranked using a voting mechanism and source stock ranked top improves the stock prediction accuracy. Finally, the prediction model was trained in the sentiment feature space using both source and target stocks, and plugged in to the prediction framework to do predictions. Support Vector Machine was used for classification and they have used the data from Hong Kong stock Exchange stocks for five years approximately. They have concluded that results based on majority voting mechanism principle shows better consistent performance than that of based on single principle stock selection.

D. Forecasting Stock Trends using Long-Short Term Memory (LSTM) Neural Network model, Multi-source Multiple Instance model:

A stock prediction model which makes use of Long-Short Term Memory (LSTM) neural network with 4 layers and 30 nodes, was suggested by Li et al.(2017) [6]. They have collected 18 million posts from stocks of CSI 200 Index for a period of around 1263 trading days and used Naive Bayes Classifier to perform classification on posts in to positive, negative and neutral. They have concluded that the prediction accuracy was 87.86 % considering the entire training data.

The consistencies among different data sources were exploited to improve the stock index movements Xi et al.[2018 [4]]. They have developed a Multi-source Multiple Instance (M-MI) model to combine the events, sentiments and the data. They have performed the experiment with five prediction models Support Vector Machine(SVM), Tensor Based Learning Approach (TeSIA), nested Multi-Instance Learning (nMIL), Open IE Multiple Instance Learning (O- MI), Without RBM Multiple Instance Learning Model(WoR – MI) and Without Hingle Loss Multiple Instance Learning Model (WoH – MI) using same set of instances collected for around two years (2015 – 2016). They have found that M- MI outperformed and SVM performed worst. They have concluded that exploiting heterogeneous information such as events, sentiments and historical features provides a better prediction accuracy considering the consistencies among them, than that of exploiting a single data source

E. Deep Learning approaches for Stock Prediction:

The performance of forecasting stock trends movement can be improved by adopting Long-Short Term (LSTM) that incorporates market trends and investor sentiments Jiahong et al[2017] [3]. They have extracted



VI.CONCLUSION

The process of predicting stock trends is difficult as the stock trends are volatile in nature and also it is affected by several information like moods of investors, economic factors etc., This study concludes that sentiment analysis can be used to assist the stock forecasting process .It also shows that most of the research that used social media information for stock prediction, used the information as a factor for short term prediction. Different social media have been used for conducting the research across different countries like Weibo (China), Twitter etc. Most of the research used Natural Language Processing Techniques for sentiment extraction and classification. Also it is clear that use of Deep Learning approaches can outperform the Conventional Machine Learning methods in predicting the stock trends but there is only few work concentrated on using Deep Learning for prediction. Table 1 gives an summary on comparative study of the different work carried out with information on Prediction Models, Social Network / Data, Classifiers used, Findings and Accuracy of prediction. It is concluded that with the help of Deep Learning approaches, the stock prediction process can be performed through the development of a specific Deep Learning Network by integrating neurons into several stages, to extract information from multiple sources like news articles, sentiments and finance indicators etc [16].

REFERENCES:

- [1] W.Long, Z. Lu and L. Cui, Deep learning-based feature engineering for stock price movement prediction, Knowledge- Based Systems (2018)
- [2] Dang Lein Minh, Abolghasem Sadeghi Niaraki , Huynh Duc Huy, Kyungbok Min and Hyeonjoon Moon,Deep Learning Approach For Short-Term Stock Trends Prediction Based On Two-Stream Gated Recurrent Unit Network (2018)
- [3] Jiahong Li,Hui Bu,Junjie Wu,"Sentiment-aware stock market prediction:A deep Learning method"14th International Conference on Service Systems and Service Management,June 2017
- [4] X. Zhang, S. Qu, J. Huang, B. Fang and P. Yu, "Stock Market Prediction via Multi-Source Multiple Instance Learning," in *IEEE Access*, vol. 6, pp. 50720-50728, 2018.
- [5] Xiaodong Li, Haoran Xie, Raymond Y. K. Lau, Tak- Lam Wong, Fu-Lee Wang, "Stock Prediction via Sentimental Transfer Learning", *Access IEEE*, vol. 6, pp. 73110-73118, 2018.
- [6] Jiahong Li, Hui Bu and Junjie Wu, "Sentiment-aware stock market prediction: A deep learning method," *2017 International Conference on Service Systems and Service Management*, Dalian, 2017, pp. 1-6.
- [7] Oliveira, Nuno et al. "Stock market sentiment lexicon acquisition using microblogging data and statistical measures." *Decision Support Systems* 85 (2016):62-73.
- [8] L. Zhang, L. Zhang, K. Xiao and Q. Liu, "Forecasting price shocks with social attention and sentiment analysis," *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*
- [9] Qian Li, Bing Zhou and Qingzhong Liu, "Can twitter posts predict stock behavior?: A study of stock market with twitter social emotion," *2016 IEEE International Conference on Cloud Computing and Big Data Analysis (ICCCBDA)*, Chengdu, 2016, pp. 359-364.
- [10] Nguyen T.H and K.Shirai, 2015 ,Topic Modeling Based Sentiment Analysis on social media for stock market prediction .Proceedings of the 7th International Joint Conference on Natural Language Processing and 53rd Annual Meeting on Association for Computational Linguistics Vol.1, July 26 – 31, 2015, Association for Computational Linguistics,Vancouver,Canada,pp:154- 1364
- [11] Nofer, M. & Hinz, O. *Bus Inf Syst Eng* (2015) 57:229.
- [12] Y. E. Cakra and B. Distiawan Trisedya, "Stock price prediction using linear regression based on sentiment analysis," *2015 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, Depok, 2015, pp. 147-154.
- [13] C. Chen, W. Dongxing, H. Chunyan and Y. Xiaojie, "Exploiting Social Media for Stock Market Prediction with Factorization Machine," *2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)*, Warsaw, 2014, pp. 142-149.
- [14] F. Xu and V. Keelj, "Collective Sentiment Mining of Microblogs in 24-Hour Stock Price Movement Prediction," *2014 IEEE 16th Conference on Business Informatics*, Geneva, 2014, pp. 60-67.
- [15] Porshnev .A,I.Redkin and A.Shevchenko,2013.Machine Learning in prediction of stock market indicators based on historical data and data from Twitter Sentiment Analysis. Proceedings of the 2013 IEEE 13TH International Conference on Data Mining Workshops (ICDMW,2013),December 7-10,2013 IEEE,Dallas,Texas,pp:440 – 444
- [16] Z. Chen and X. Du, "Study of Stock Prediction Based on Social Network," *2013 International Conference on Social Computing*, Alexandria, VA, 2013, pp. 913-916.
- [17] Al Augby,S.H.,2015 Text Mining method in evaluation of media's impact on market value ratios.Ph.D thesis,University of Szczecin ,Szczecin ,Poland
- [18] D.Duong ,T.Nguyen,M.Dang, "Stock Market Prediction using Financial News Articles on Ho Chi Minh Stock Exchange"- Proceedings of 10th International Conference on Ubiquitous Information management and communication, 2016 ,Article No 71.
- [19] Al-Augby,Noor Al-Musawi and Abdul Hussein Mezher,Stock Market Prediction Using Sentiment Analysis Based on Social Network:Analytical Study,Journal of Engineering and Applied Sciences,2018,pp:2388 - 2402