



SENTIMENT ANALYSIS FOR ONLINE STOCK MARKET NEWS USING RSS FEEDS

SV.Shri Bharathi¹, Angelina Geetha²

Department of Computer Science and Engineering, B.S.Abdur Rahman University, Vandalur, Chennai, Tamil Nadu, India.

Abstract

The most important factors which influence the stock market are the stock news. This research is an attempt to build a model that predicts news polarity which may affect changes in stock price movement trends. The novelty of our approach is efficient prediction of model that scores emotions from all relevant real time stock news available in public domain. Our method use the sentiment analysis in Really Simple Syndication (RSS) news feeds that has an impact on stock market values. Hence stock market RSS news feed data is collected for a period of time. The proposed paper uses the algorithm for sentiment analysis to establish the sentiments in RSS news feeds. In our experimental study, the algorithm method reveals that the stock news sentiment can be used in predicting stock price fluctuations, whether up or down.

Index Terms: Sentiment Analysis, Opinion Mining, RSS Reader, Stock Market News.

I. INTRODUCTION

Opinion Mining (OM) is a recent sub-discipline of text classification which is concerned not with the topic, a document is about, but with the opinion it expresses. Opinion mining is also called sentiment classification, sentiment analysis, opinion mining for text or sentiment orientation analysis. An interesting area to explore is opinion analysis of news document. Sentiment analysis is often used in opinion mining to identify sentiment, affect, subjectivity, and other emotional states in online text.

In the finance field, stock market and its trends are extremely volatile in nature. It attracts researchers to capture the volatility and predicting its next moves. Investors and market analysts study the market behaviour and plan their buy or sell strategies accordingly. As stock market produces large amount of data every day, it is very difficult for an individual to consider all the current and past information for predicting future trend of a stock. Mainly there are two methods for forecasting market trends. One is technical factor and other is fundamental factor.

Technical analysis considers past price and volume to predict the future trend. In the same way fundamental analysis considers the fundamental analysis of a business involves analysing its financial data to get some insights.

This research follows the fundamental factor technique to discover future trend of a stock by considering news articles about a company as prime information and tries to classify news as good (positive) and bad (negative). If the stock news sentiment is positive, there are more chances that the stock price will go up and if the news sentiment is negative, then stock price may go down.

This paper is organized as follows: Section 2 surveys the related work. Section 3 discusses the Opinion Analysis and RSS News Feeds. Section 4 describes the Stock Market news. Section 5 describes System for Sentiment Analysis for Online Stock market news using RSS Feeds. Section 6 presents the experimental results. Finally, Section 7 gives the conclusion of work.

II. RELATED WORK

In recent years, significant efforts have been put into developing models that can predict the

future trend of a specific stock or overall market. Most of the existing techniques make use of the technical indicators. The prediction of Chinese stock market with the help of public moods is extracted from micro blog feeds [1].

Karamibekr work focuses on the subjectivity analyses of social issues [2]. The subjectivity of a document strongly depends on its sentences. They proposed a lexical-syntactical approach to recognize and classify subjectivity at the sentence level and consider the role of various opinion terms, especially verbs on opinions regarding social issues.

Li et al proposed a generic stock price prediction framework to enable the use of different external signals to predict stock prices [3]. Evaluations were performed at individual stock, sector index, and market index levels.

Another approach was proposed to mine Twitter data [4]. By extracting ambiguous textual tweet data through NLP techniques to define public sentiment, then make use of a data mining technique to discover patterns between public sentiment and real stock price movements. Al-radaideh et al proposed a decision tree classifier which is one of the data mining techniques that helps to take the decision in the stock market [5].

The proposed approach utilizes the results of a lexicon emotion analysis conducted on crowd-annotated news to extract various types of public emotions from daily news articles. Another approach was a model to score news articles regarding public emotions, and to identify which news sections and emotions cause movements in a stock market index [6].

To deal with the uncertainty factors, a fuzzy logic approach based on grid partition is adopted in this paper while predicting the stock price of any company. Also the premise and consequent parameters of the learning rules are optimized in an adaptive fashion using a hybrid neural learning mechanism [7].

This works of Cheng et al focused on the Joint Aspect/Sentiment (JAS) model, to jointly extract aspects and aspect-dependent sentiment lexicons from online customer reviews [8]. It then applies the extracted aspect-dependent sentiment lexicons to a series of aspect-level opinion mining tasks, including implicit aspect identification, aspect-based extractive opinion summarization, and aspect-level sentiment

classification.

A novel stock selection model with discrete and continuous variables algorithm is introduced for model optimization [9]. Different textual representations of news articles were examined to predict future stock price, which was compared to linear regression with SVM [10].

III. OPINION ANALYSIS AND RSS FEEDS

Opinion mining or sentiment analysis refers to activities that automatically identify a sentiment (such as a positive or negative sentiment) from a group of words such as a sentence or a document. Newspapers generally attempt to present the news objectively, but textual affect analysis in news documents shows that many words carry positive or negative emotional charge.

Sentiment classification could be done in word/phrase level, sentence level and document level. Sentiment analysis has now become the dominant approach used for extracting sentiment and appraisals from online sources. Subjectivity analysis focuses on dividing language units into two categories: objective and subjective, whereas sentiment analysis attempts to divide the language units into three categories; negative, positive and neutral.

Three levels of sentiment analysis are,

Document level: The whole opinion of the document is identified as positive, negative or neutral.

Sentence level: Identify if a sentence is opinionated and whether the opinion is positive, negative or neutral.

Entity and Aspect level: Extract the object attribute that are the subject of an opinion and the opinion orientations.

This research work explores the task of automatic identification of news opinions with the help of RSS Feeds. From the web pages, RSS feed reader reads the required news content. This RSS feed helps to collect the stock market News as a Dataset.

Really Simple Syndication (RSS) is a format for delivering regularly, the ever changing Web content. Many news-related sites, Weblogs and other online publishers syndicate their content as an RSS feed to whoever wants it. RSS mostly, uses XML to deliver updated content on the Web. RSS news feeds are widely and easily available to the users from various news sites without visiting the sites individually.

RSS feed reader reads the required news content from the web pages e.g. title, description, date, author, link etc. in the format of XML. With RSS feed, users can finally separate wanted information from unwanted information. If the RSS news feed is positive then this tends to have a positive effect on stock markets. On the other hand, if the RSS feed is negative then this tends to have a negative impact on stock markets and makes the stock market values go down.

IV. STOCK MARKET NEWS

Stock market prediction is one of the most difficult tasks in the financial markets which are influenced by many external social-psychological and economic factors.

A. Impact of RSS News Feed on Stock Market

Market and stock exchange news are special messages containing mainly economic and political information. Some of them are carrying information that is important for stock market prediction. There are various types of financial information sources on the Web which provide

electronic versions of their daily issues. All these information sources contain global and regional economic news, citations from influential bankers, as well as recommendations from financial analysts.

They have shown that economic news always has a positive or negative effect on the number of traded stock. They used salient economic news as proxy for public information. They have found that news have impact on measures of trading activity including return volatility, price volatility, number of shares traded, and trading frequency. Sentiment polarity news provides an efficient result to the stock market forecasters when to buy or sell their stocks.

V. PROPOSED WORK

The system automatically identifies the news opinions with the help of RSS news feeds and predicts the stock market movement whether rises up or down. The Fig. 1 describes system for Opinion Mining on Stock market News Prediction using RSS News feed.

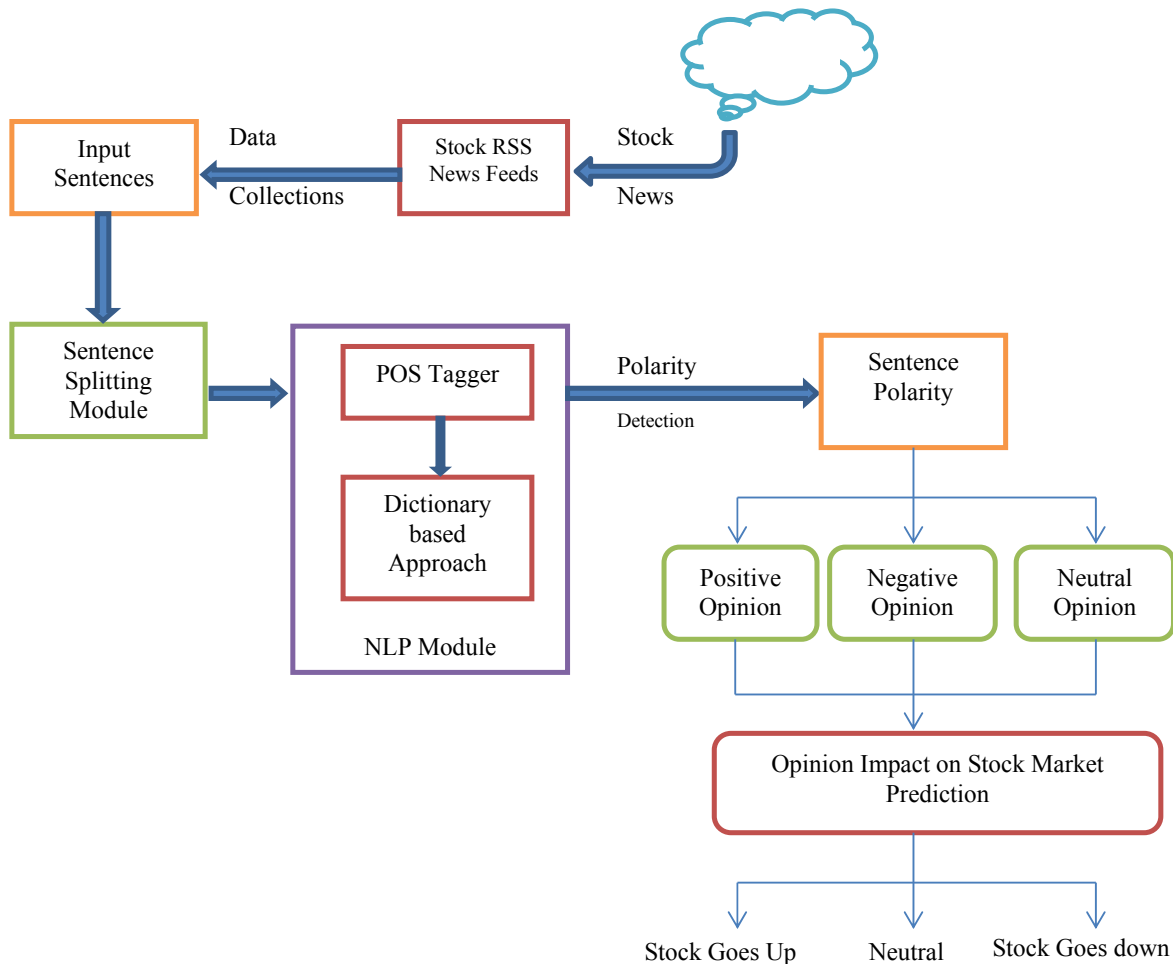


Figure 1. System for Sentiment Analysis for Online Stock market news using RSS Feeds

A. Stock RSS Feeds

From the relevant web sites, RSS feeder reads the required content such as title, description etc. in the form of XML. These feeds are collected inside this module.

B. Input Sentences

All the collected RSS stock news feeds are stored inside the input sentence module as a whole document.

C. Sentence Splitting Module

The sentence splitting module is the one which cleans the news data and splits it into parsed sentences. A text document is used to collect all the parsed news data for the testing purpose. The document contains the RSS news data in the form of sentence by sentence

D. NLP Module

Natural language processing (NLP) is used to identify and extract subjective information in source materials. In order to identify the new opportunities and to maintain the reputations, business people usually view the reviews, ratings, recommendations and other forms of online opinion.

This module is used to determine sentiment for the words having a positive/negative/neutral sentiment. Here the part-of-speech tagger, dictionary based approach are used to find the polarity of the sentence, to conclude whether the sentence is positive, negative and neutral.

Part-of- Speech Tagger: A Part-Of-Speech (POS) tagger is also known as grammatical tagging or word-category disambiguation. It is a piece of software that reads text in some language and assigns parts of speech to each word, such as noun, verb, adjective, etc.,

Dictionary based Approach: The dictionary is used to find the opinion words and their polarities. Dictionary-based techniques use synonyms, antonyms and hierarchies in WordNet (or other lexicons with sentiment information) to determine word sentiments. Three sentiment numerical scores used are: Obj(s), Pos(s), and Neg(s) and describing how Objective, Positive and Negative synset are.

E. Sentence Polarity

Each sentence polarity value is calculated in this module. In general, the score ranges from

0.0 to 1.0 and their sum is 1.0 for each synset. If the polarity value of that sentence is positive then that sentence is considered as a positive sentence. If polarity value is negative then it is considered as negative sentence. If it is 0.0, then it is considered as a neutral sentence.

Equation (1) is used to calculate the Sequence of words.

$$\text{Sequence of words (W)} = W_1 + W_2 + \dots + W_n \quad (1)$$

$S(W_{+ve})$ = Set of Positive Sentiment words.

$S(W_{-ve})$ = Set of Negative Sentiment words.

n = Number of words.

F. Opinion Impact on Stock Market Prediction

After identifying the opinionated score of each sentence, the polarity score value is classified as positive opinion, negative opinion and neutral. If the opinion is positive then the stock goes up and if the opinion is negative the stock goes down. Based on the results the stock market trend goes up, down or neutral.

VI. EXPERIMENTAL RESULTS

In the experimental study the stock market predicting is collected for the company ARBK from Amman Stock Exchange (ASE). The sentiment RSS news feed for Arab Bank (ARBK) Company (5) is collected from <http://investing.einnews.com/news/ase-stock> for the period of 2005 to 2007. The polarity value is calculated for specific month of 499 sentences. Now the sentiment prediction result shows that it is efficient than previous one.

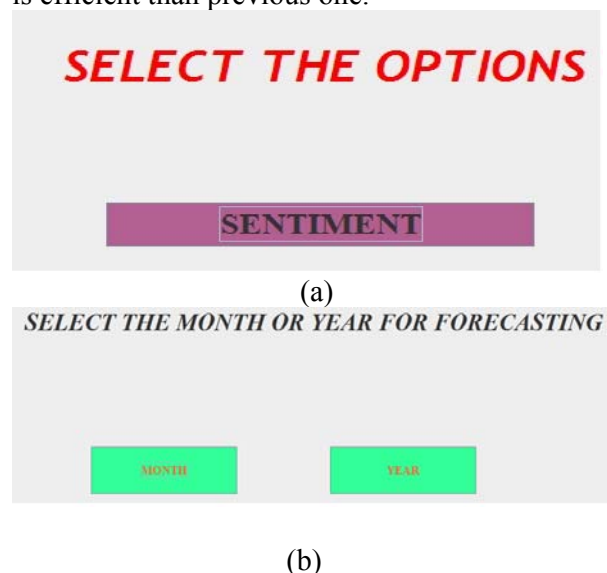


Figure 2. The selection of month and year for opinion analysis (a), (b)

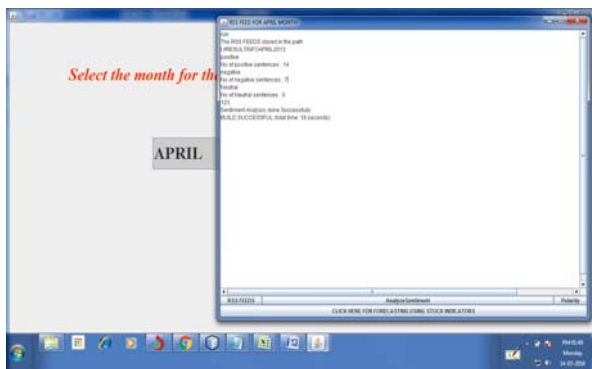


Figure 3. Calculation of Over all Positive and negative scores.

The selection of sentiment analysis for the specific month is shown in Fig.2 (a), (b). Also the overall positive and negative polarity scores are shown in Fig. 3. The below Table I shows the opinion score value for the month of April 2007. From this graph the positive score value is greater than negative one. This result shows that the stock value goes up for that month. The Left column shows the number of sentences and the bottom value shows the date of the month. Highest line score is for positive value and second line is for negative score value.

ARB K Comp any	ID3 (Previous)			Sentiment Analysis(Propose d)		
	Pos s	Ne g	Neut ral	Pos	Ne g	Neut ral
Total Instan ces (499)	20 0	17 5	19	277	20 0	22
Correc tly Classif ied	12 0	11 2	12	248	16 8	18
Precisi on %	60 %	64 %	63%	89.5 %	84 %	81.8 %

Table I Sentiment score value for the month of April 2007.

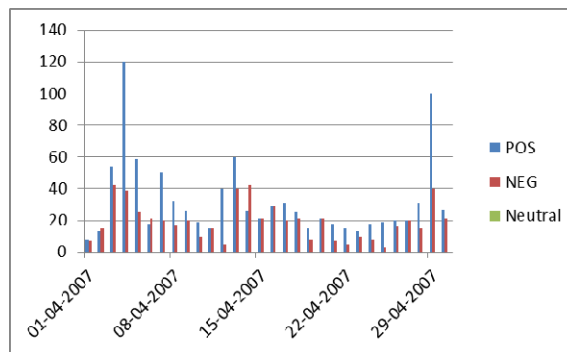


Table 2. Accuracy Classification for sentiment analysis and ID3 method.

The above Table 2 shows the accuracy prediction of the proposed method.

Accuracy refers to the closeness of a measured value to a standard or known value. Precision refers to the closeness of two or more measurements to each other. Equation (2) is used to calculate the precision score values.

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP}) \quad (2)$$

Compared with the above tabular value the proposed system produces better than the previous one.

VII. CONCLUSION

This paper builds a predictive model to predict sentiment around stock news. First the relevant real time RSS stock news has been filtered from Arab Bank (ARBK) Company, and then they have been analyzed to predict the sentiments score values whether it is positive, negative or neutral. Such proposed model can be a helpful tool for the investors to take the right decision regarding their stocks. Finally sentiment polarity news provides an efficient result to the stock marketers when to buy or sell their stocks. The future of the system focuses on the evaluating the impact of using negation and valence shifters in addition with sentiment news which may improve the accuracy.

REFERENCES

- [1] D. Yan, G. Zhou, X. Zhao, Y. Tian and F. Yang, "Predicting Stock using Micro blog Moods," Journal on China Communications, vol. 13, no. 10, pp.244- 257, 2016.
- [2] M. Karamibekr, A. A. Ghorbani, "Sentence Subjectivity Analysis in Social Domains,"

- In: Proc. of IEEE/WIC/ACM International Conf. On Web Intelligence (WI) and Intelligent Agent Technology (IAT), pp.268-275, 2013.
- [3] X. Li, H. Xie, Y. Song, S. Zhu, and Q. Li and F. L. Wang, "Does Summarization Help Stock Prediction? A News Impact Analysis," IEEE Intelli. Systems, vol. 30, no. 3, pp. 26-34, Jan 2015.
- [4] B. Li, K. CC Chan and C. Ou, "Public sentiment analysis in Twitter data for prediction of a company's stock price movements," IEEE 11th Intl. Conf. On e-Business Engineering (ICEBE), pp. 232-239, 2014.
- [5] Q. A. Al-radaideh, A. A. Assaf, E.Alnagi, "Predicting Stock Prices using Data Mining Techniques," In: Proc. of International Arab Conf. on Information Technology, pp. 1-8, 2013.
- [6] C. Wong; I. KO , "Predictive Power of Public Emotions as Extracted from Daily News Articles on the Movements of Stock Market Indices," IEEE/WIC/ACM Intl. Conf. On Web Intelligence (WI), Oct 2016.
- [7] A.Chakraborty, D.Mukherjee, A.Dutta, D.K.Kole, "An Effective Stock Price Prediction Technique Using Hybrid Adaptive Neuro Fuzzy Inference System Based on Grid Partitioning," Journal on Advances in Intelligent Systems and Computing, vol 385. Springer, 2016.
- [8] X. Xu, X.Cheng , S. Tan ,Y. Liu , H. Sh, "Aspect-level opinion mining of online customer reviews," China Communications. Vol. 10, no.3, pp. 25-41, 2013.
- [9] L. Yu, L. Hu and L. Tang, "Stock Selection with a Novel Sigmoid-Based Mixed Discrete-Continuous Differential Evolution Algorithm", IEEE Transactions on Knowledge and Data Engineering, vol. 28, no. 7, pp.1891- 1904,2016.
- [10]P.Robert, Schumaker, and H.Chen, "Textual Analysis of Stock Market Prediction Using Breaking Financial News: The AZF in Text System", In: Proc. of 12th American Conf. on Information Systems, 2009.