Sentiment Analysis in Financial Markets

A Framework to utilize the Human Ability of Word Association for analyzing Stock Market News Reports

Patrick Uhr, Johannes Zenkert, and Madjid Fathi

University of Siegen
Institute of Knowledge Based Systems
Germany

***Abstract*—As financial markets getting faster and more complex, it is difficult for market participants to manage the information overload. Sentiment analysis is a useful text mining method to process textual content and filter the results with analysis methods to relevant and meaningful information. The paper in hand introduces a new method for sentiment analysis in financial markets which combines word associations and lexical resources. Based on stock market news from January 2000 to February 2014 we analyzed documents on different levels. The results are presented and evaluated in this paper.**

Keywords: Sentiment Analysis, Word Association, Text Mining

I. INTRODUCTION

The evolution of research works in the field of Quantitative Finance started in the early 1950s with the first computerized analysis of time series. Since then it was discussed in several theories if markets are efficient, prices reflect all known information and adjust quickly to new information, or do they just follow a random walk. Today the theories are still well-known, especially the Efficient Market Hypothesis (EMH) from Eugene Fama [1]. Three levels of market efficiency were described in this theory. The weak form, which indicates there is no possibility for prediction of the future price from the past ones. Instead, price movements are following a random-walk, also discussed earlier in [2]. The semi-strong efficiency means the present price reflects all publicly available information and only insider information, which is hidden, could lead to an advantage. The third one is the strong-form efficiency which assumes every information - public and hidden - is already represented by the current price.

What one can discover from the three levels of market efficiency is that financial markets do not fit in any of these levels. There is the prohibited practice of insider trading, which should not give any advantage in the strong form of efficiency. There are (irrational) overreactions at bull and bear markets and several repeating patterns in approximately predictable price movements [3]. In conclusion only a mixture of inefficiency and efficiency can be the right point of view.

The adjustment of prices in the semi-strong efficiency form due to new information is an extensive researched market principle. Today the role of information had never been more important. Financial institutions pay millions to benefit from technological advantages to get information in the first place. In competitive financial markets the latency of 59 milliseconds between the stock exchanges in New York and London seems to be a priceless advantage. Lately, in 2013, it was reduced by six milliseconds, down from 65 milliseconds, with a new $300 million transatlantic connection. High Frequency Trading (HFT) benefits the most from technology that increases order execution and trading speed. This development is critically reviewed in many investigations, especially the amplifying effect on prices in extreme high volatile market situations [4]. In 2012 Morgan Stanley found that over 80 percent of all market transactions are realized by high-frequency trading engines. The strategies from HFT companies range from market making, statistical arbitrage to low-latency- and news-based trading. The profitability of them was recently shown in March 2014 in the IPO Registration Form S-1 from Virtu Financial Inc. with an impressive track record of one single losing trading day in 1238 days of trading.

Besides the technological advantage of having access to information as early as possible, the intelligence to process the information is an important advantage in financial markets. Algorithms for textual analysis and news-based trading are used to determine if the current news flow about companies and their stock prices, currencies or commodities are positive or negative. In this context the application of sentiment analysis, also known as opinion mining, is used on big data with text mining methods to handle unstructured textual information.

This paper illustrates a practical application of the sentiment analysis on financial market news in German language in the period January 2000 through February 2014. We integrated several text mining methods in a software prototype for calculating results on different analysis levels. The following section gives an overview on sentiment analysis in general, and the usage in financial market applications. Afterwards our software implementation will be introduced in detail. Finally, the calculated sentiment scores are evaluated and the results are discussed in the conclusion.

II. SENTIMENT ANALYSIS

Over the last years the interest in sentiment analysis has grown rapidly. In the first approaches developed in the sentiment analysis research, terms with high occurrence numbers measured by term frequency or phrases for semantic orientation were tracked and affected the algorithms and models for document classification [5]. Later on, the level was reduced to overview segments of texts and even sentences due to the fact that documents contain multiple entities, which can express their opinions. In this case the identification of subjectivity and objectivity of sentences and their entities is an important task [6]. The finest grade for sentiment research is the entity- or aspect-level, also called feature-level [7]. At the sentence-level, part-of-speech (POS) information are used to extract sentiment polarity from nouns, verbs, adverbs, and adjectives, as well as word sense disambiguation with the help of natural language processing (NLP) techniques. Furthermore n-grams and co-occurrences of words are typical pattern-based methods for sentiment analysis. N-gram, bi-grams, tri-grams or uni-grams have been researched by [8] and [9] with contrary results on sentiment classification of movie and product reviews.

The most common methods in sentiment analysis are dictionary-based approaches based on lexical resources [10]. Sets of known positive and negative categorized words are collected to a dictionary, which is extended (automatically or manually) with synonyms and antonyms in further steps [10]. Another approach is to assign terms to emotions, such as “afraid”, “bored”, “happy”, etc. to get a sentiment value based on different categories of emotions [11]. Other lexical resources provide predefined scores for words, word families or n-grams. In addition to lexical affinity and keyword spotting, statistical methods and concept-based techniques are used in the majority of cases [12].

Problems in the sentiment analysis are still existent. Negation handling and word sense disambiguation are difficult and not solved NLP tasks [10]. Moreover opinion spam is a problem which started especially with the rise in popularity of social media [13]. It is vital to detect spam or ironical comments in order to extract true sentiment evaluation and uncover fake reviews, wrong reputation or promotions.

A. *Financial market information overload*

Market participants have recently raised growing interest in the field of sentiment analysis. With the development of algorithmic trading systems human-decision making has reached its limit and cannot keep up with processing and execution of orders and decisions. Additionally, the information overload is not manageable without support. To manage all information, analysis methods are applied on different category groups which are expected to hold information about sentiment polarity [14]. Forums, blogs and wikis are seen to be the most representative types of textual-data where users can express their opinion. Also news, research reports and finance-related content generated by firms are analyzed [14].

B. *Applying the associaton concept CIMAWA for Sentiment Analysis*

An innovative method to create a sentiment evaluation is to combine two concepts, namely keyword spotting and word association. The Concept for the Imitation of the Mental Ability of Word Association (CIMAWA) is developed at the Institute of Knowledge Based Systems to measure the association between words [15]. It was designed to simulate the Human Word Association (HWA), based on large collections of text documents.

$$CIMAWA_{ws}^{\zeta}(x(y)) = \frac{Cooc_{ws}(x,y)}{(frequency(y))^{\alpha}} + \zeta * \frac{Cooc_{ws}(x,y)}{(frequency(x))^{\alpha}} \quad (1)$$

The equation (1) describes the strength of association between two words ‘x’ and ‘y’ by taking into account the co-occurrences ($Cooc_{ws}(x,y)$) in a predefined text window (ws). The damping factor ζ is used to balance the first asymmetric and second symmetric part of the hybrid association formula [15]. By solving the equation for different word combinations one can create all associations (represented, scaled and ordered by the numeric CIMAWA values) from a selected document corpus.

As a measure for sentiment analysis in a predefined time period the concept and equation of CIMAWA was adapted to be more flexible to changes of word associations over the time:

$$CIMAWA_{\mathrm{T},ws}^{\zeta}(x(y_i)) = \frac{\sum_{t=1}^{\mathrm{T}}\left[\left(\frac{Cooc_{ws}(x,y_i)}{\left(frequency(y_i)\right)^{\alpha}}_t + \zeta * \frac{Cooc_{ws}(x,y_i)}{(frequency(x))^{\alpha}}_t\right) * \theta_{y_i}\right]}{n} \quad (2)$$

All co-occurrences ($Cooc_{ws}(x,y_i)$) have a subindex $i = 1, \ldots, n$ at keyword ‘y_i’ for an increment operation. Word ‘x’ is a fixed term in the calculation process and will not be changed. The increment operation only counts if ‘y_i’ (inside a text window size (ws)) is present in a resource (Θ) (e.g. a wordlist), with evaluated entries (θ) and finishes at n. For the damping factor ζ values in the interval [0.4 – 0.6] achieved best results in previous studies [16]. For this application damping factor $\zeta = 0.5$ is used.

The formula is applied in temporary changing content. With timestamp information (t), all sub periods (starting with $t = 1$), combined in a maximum time window (T) are observed. All calculated CIMAWA associations are multiplied by the valuated wordlist scale Θ, with value entry θ_{y_i} if keyword ‘y_i’ is included. Finally, the sum of all multiplication results is divided by the last increment value n.

To create a dynamic sentiment value - similar to a moving average - this calculation is done for each document of a document corpus with timestamp information. The time window (T) is moving forward and reviews positive and negative associations in small parts of the whole corpus. The advantage of this extension is the in-depth historical look-back over the time and determination of the changing strength of word associations in combination with the fast processing concept of spotting relevant keywords to filter results. Fast or slow changing sentiment values can be easily created by adjustment of the time window.

III. DEVELOPING A SENTIMENT ANALYSIS APPLICATION

An implementation of several text mining techniques for sentiment analysis on financial market news was realized by developing a software application. The application was programmed in C# language using the Model View ViewModel (MVVM) pattern and a Windows Presentation Foundation (WPF) User Interface. The database connection is supplied by the ADO.NET Entity Framework. It connects the application to the relational MySQL database and maps the contained tables and relations between them to objects and properties.

A. Concept

The concept for the development of the sentiment analysis application is shown in Fig. 1. The workflow starts by setting up a database, which holds all the information for the application.

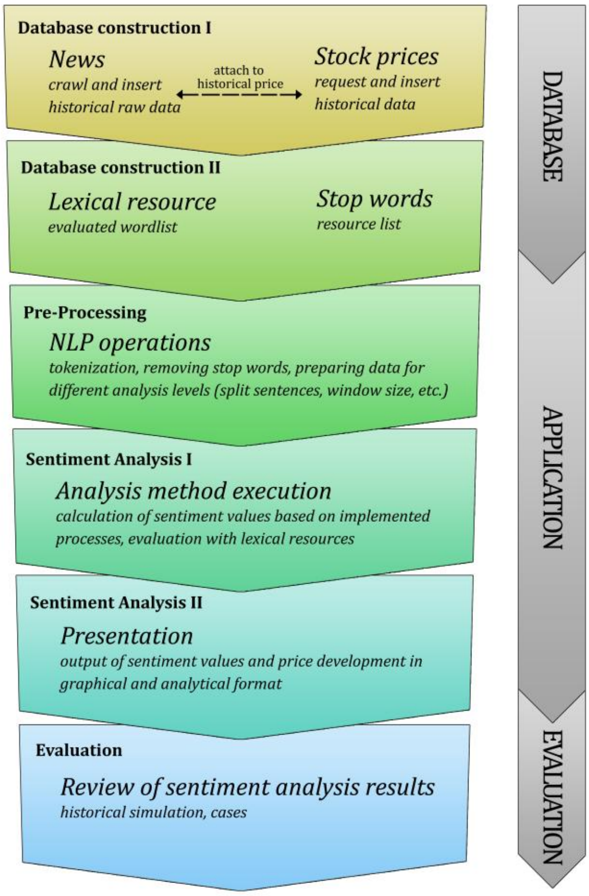

Fig. 1. Concept for the sentiment analysis software application and its workflow

A news corpus with 918,427 documents was collected from finance-related content with news, researches, ticker news, analyst ratings, recommendations, comments, etc. from July 2000 to February 2014. Each news consists of timestamp, headline, message, category and frequent attached stock information. Stock quotations on a daily basis in Open, High, Low, Close format (OHLC) over the same time period were added to the database and edited. The assignment of news to stock prices and all import operations into the database were done with language scripts in the software environment for statistical computing and graphics R.

News messages and headlines are in an unstructured data format, which cannot be processed and analyzed without preprocessing. Documents have to be prepared with NLP techniques to enable the analyzing machine to get a basic understanding of the textual information. Typical tasks in the NLP pre-processing are tokenization and removing stop words [17]. At the process of tokenization every element of the text is separated at punctuation or whitespace to represent a token. Stop words are common non-meaningful tokens, which do not provide any or only little information [17]. To speed up processing time of analysis methods they are removed. The sentiment analysis operates on different document levels. For that reason the pre-processing considers to split text into sentences and subsets of tokens from sentences, as well as window size separation of tokens for detailed analysis.

After pre-processing the actual sentiment value calculation can be executed with lists of tokens and an evaluated wordlist. For German language the SentimentWortschatz (SentiWS) - a publicly available resource for sentiment analysis, developed and introduced in [18], was chosen for the application.

B. Sentiment scoring algorithms

Sentiment values are observed on different levels. The analysis is refined from the whole document with all information and multiple entities to a more detailed stock specific search on sentence and window size level.

1) Document: The document level sentiment is calculated by summing up all values from matching wordlist entries and the pre-processed token list. For comparability reasons the sum is divided by number of tokens.

2) Sentence: The sentence level sentiment is calculated in a similar way. However the token list consists only terms from relevant sentences, identified by an alias for stock corresponding company names. All values from the matching wordlist entries are summed up and divided by the relevant sentence token amount to represent a sentiment value on sentence level for every news. Furthermore, the CIMAWA sentiment calculation is done on sentence level. Fast sentiment changes are tracked by a 1-week moving time window (T). Slower changes in strength of word associations are recognised by a 3-month version. A sentiment value based on word associations with CIMAWA is calculated for every document on sentence level with all tokens from relevant sentences identified by the company alias.

3) Window Size: Additionally, the same calculation is done with a fixed text window size ws of 10. The algorithm is searching for the relevant company alias and collects five tokens on the left and five tokens on the right side. Based on the token list the CIMAWA calulation is done for a 1-week and 3-months version.

4) N-Gram: Analyst commentary and recommendations are very similar from time to time. With price targets and ratings like “buy”, “hold”, “sell”, “out- and underperform”, “over- and underweight”, etc. the n-gram approach seems to be promising. A custom-weighted wordlist with 93 entries was created to track all major possible up- and downgrades from different levels (e.g. “from buy to hold”, “from sell to neutral”, “from hold to strong buy”) for stocks within the application. The algorithm searches for the worldlist n-grams on sentence level, identified by the company alias.

C. Application

The application was built with three major components. Firstly, the news component to read, edit, delete and insert news. It also shows and calculates document level sentiment values. Secondly, in the stock component the user is able to look up stock quotations, company profiles, graphical chart visualization with prices, sentiment indicators and to load all attached news to a selected stock. Thirdly, the application provides an analysis component with many graphical and analytical possibilities to explore sentiment calculations. Fig. 2 shows the graphical user interface of the analysis component with a tab-oriented design for different analysis methods.

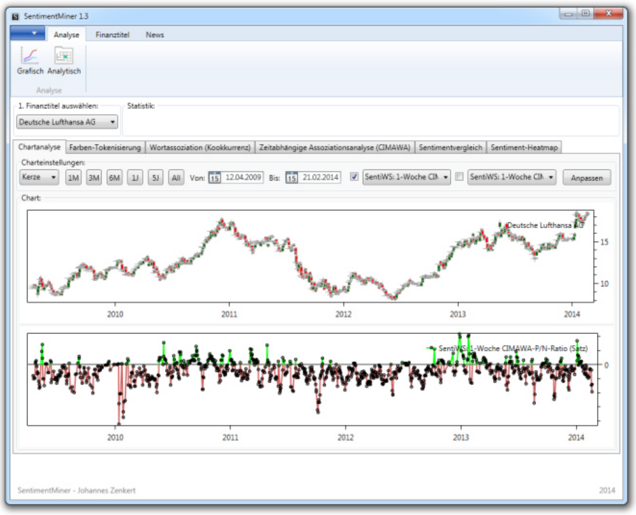

Fig. 2. Analysis component of the sentiment analysis application, stock price from “Deutsche Lufthansa AG” [from Apr. 2009 until Feb. 2014] with 1-week CIMAWA-based sentiment calculation (on sentence level)

IV. Evaluation

For evaluation several stock prices have been screened and compared to the sentiment measures calculated by the application. In the following paragraphs a small part of the results is presented. First, the application’s graphical sentiment analysis with the concept of CIMAWA and a look at analyst recommendations. Afterwards, the document and sentence level sentiment is consolidated to monthly data to evaluate sentiment changes in a long term perspective.

A. CIMAWA-based sentiment comparison

The CIMAWA-based sentiment was evaluated by comparison of stock prices and sentiment development. In Fig. 3 the stock prices of Cisco Systems Inc. and Amazon.com Inc. are evaluated by the CIMAWA-based research on their corresponding news. For Cisco Systems in the upper part of Fig. 3 the sentiment indicator is very volatile and the time window period of one week in the extended CIMAWA calculation reacts fast to changing positive or negative word associations. The calculation is done on seven days, which means weekend news data is included. On the user interface in the analytical component of the application one can see the development in price and sentiment over the time. The line in the graph is interpolated, which means, with more news it becomes more accurate.

Fig. 3. Top: Cisco Systems Inc. [Jan. 2013 to Feb. 2014] with 1-week CIMAWA-based sentiment (on sentence level), Bottom: Amazon.com Inc. [Jan. 2013 to Feb. 2014] with 3-month CIMAWA-based sentiment (on window size level)

In the lower part of Fig. 3 Amazon.com Inc. is described by the 3-month version of the CIMAWA-based sentiment with a text window size level (*ws* 10). By measuring the obvious visible correlation one can see parallel top and bottom development in price and sentiment over the time. The value of the sentiment does not necessarily need to be positive at all to indicate a rising price. An uptrend in negative sentiment, can also lead to an increasing price. As well as positive sentiment may is reduced in a downtrend with a decreasing price reaction. The idea of mean reversion of extreme values in stock returns, one of the earliest research was done by [19], which says prices tend to return to a mean value after an overreaction, can also be applied and seen in sentiment scales.

Fig. 4 shows local extreme values of the daily Adidas AG price and attached CIMAWA-based sentiment in the time from Oct. 2010 through Feb. 2014. The chart was edited with green and red circles to highlight these values. Nearly all highs and lows of the price time series were captured by the sentiment scale based on sentence and window size level successfully.

Fig. 4. Adidas AG [Oct. 2010 to Feb. 2014] with 3-month CIMAWA-based sentiment (on window size and sentence level)

B. *Analyst Recommendations*

In Fig. 5, the major up- and downgrades from Apple Inc. are visualized. Based on the n-gram sentiment analysis of corresponding news on sentence level, all positive and negative recommendations are tracked and presented in the user interface.

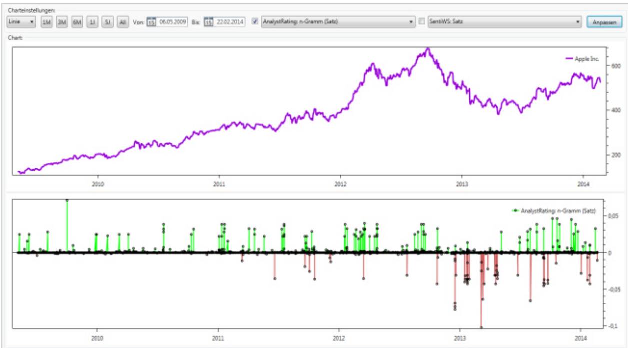

Fig. 5. Up- and Downgrades from Apple Inc. [May 2005 to Feb. 2014] measured by the application's AnalystRating sentiment (on sentence level)

Interestingly, the analyst commentary and recommendation is often clustered over the time. It appears to be like a momentum effect, if some financial institution started to up- or downgrade other analysts tend to do the same, especially after earnings conferences and announcements. In the example of Apple in Fig. 5, the stock price seems to be also driven by these clustered analyst up- and downgrades as well as recommendations (Fig. 5 green and red bars) in the analyzed time period.

C. *Equal-weight portfolio vs. Document- and sentence-level Portfolio*

The sentiment analysis is often seen as a tool to predict direct changes and movement in stock prices due to immediate sentiment changes. Many research work has been done by several authors in the field of prediction a stocks next move. On the one hand the focus lies upon price reactions and pattern research in time series analysis. Machine learning techniques are investigated in [20], whereas on the other hand [21] focuses on a textual approach and achieved 87.6% accuracy on predicting the daily up and down movement of the Dow Jones Industrial Average.

With the results of the application's calculations it turned out, that sentiment changes are as noisy as stock prices in a short-term time frame, which complicates the prediction a lot. On contrast, a long-term sentiment, like the CIMAWA-based sentiment, could be useful to indicate constant market conditions. For that reason a portfolio of stocks with different long term perspectives was evaluated on a monthly time frame with the results of the application's news document and sentence sentiment analysis in *R*.

The equal-weight portfolio consists of 30 stocks, all held during all up- and down-phases of the market from Nov. 2000 to Feb. 2014. The 30 stocks were randomly selected during five simulations from a universe consisting the components of the DAX, Dow Jones Industrial Average and well known titles from Nasdaq 100. Blue chips were chosen because of the higher news amount in the news corpus to generate results based on more data. The performance of the five simulated equal-weight portfolios can be seen in Fig. 6 in black color. The mixture of different performing stocks makes the equal-weight portfolio familiar to the whole market development. Price effects and divergences from different currencies were not taken into account, only the performance was measured as relevant characteristic by the monthly rate of change.

Fig. 6. Portfolio simulation results with monthly data. Top: Sentiment-based portfolio on document level, Bottom: Sentiment-based portfolio on sentence level

For the document- and sentence-level portfolios a rebalance of investment in stocks was considered at the end of every month by investigating the average of all sentiment values. So every stock in the portfolio stock universe was observed on its monthly news. Positive news lead to an investment. In this logic negative news indicate a decreasing stock price.

As a trading approach for the document- and sentence-portfolios two exponential moving averages (EMA) are used to track the consolidated monthly sentiment values. One of them is a “faster” 3-month period EMA which reacts faster to long-term sentiment changes than the “slower” 10-month period EMA. The reason to prefer the exponential version of the moving average over a smooth or simple one is the higher influence of latest monthly sentiment values. A buy signal for stocks is triggered if the fast reacting 3-month EMA crosses or notes over the slower 10-month EMA at the end of the month. Stocks are sold in the contrary scenario. For simulation reasons both portfolios were composed by long (buy only) and long/short positions (with speculation for decreasing prices) in every simulation. In the end all ten document- and ten sentence-portfolios are equal- weighted into the two balanced ones.

As a result, shown in Fig. 6, the document level (red) in the upper part of the figure performed very well in downturns of the market, the sentence level (green) in the lower part performed better in rally phases. The implication of the outperformance on document level against the sentence level could be a better risk measure by the broader analysis on document level which leads to fewer investments by presentiment of falling prices before the market plunges.

V. CONCLUSION AND FUTURE WORK

Sentiment analysis on financial market news provides meaningful information on different time scales for decision making and risk management. Stock and news corresponding sentiment values tend to be noisy in smaller timeframes which makes prediction a difficult task. The analysis with the help of the developed application provided an insight into sentiment analysis on larger time scales. The CIMAWA-based word association measurement and the evaluation with a lexical resource for sentiment analysis shows good results. A comparison of the CIMAWA-based portfolio to the presented document- and sentence-level portfolios may show the improvement in sentiment analysis on financial market news with word associations and the potential of this approach against the conventional concepts. The application will be updated in the future to cover intraday market information and sentiment changes. Especially, live updating social media and up-to-the minute news for real-time sentiment analysis on new content and topics will be interesting research area. Furthermore we plan to integrate evaluation procedures to recognize repeating and prominent patterns in the database.

REFERENCES

[1] E. Fama, “The Behavior of Stock Market Prices,” *Journal of Business*, vol.38, pp.34-105, 1965.

[2] M. G. Kendall, “The Analysis of Economic Time-Series-Part I: Prices,” *Journal of the Royal Statistical Society*, vol.116, no.1, pp.11-34, 1953.

[3] G. Caginalp and H. Laurent, “The predictive power of price patterns,” *Applied Mathematical Finance* 5, pp.181–205, 1998.

[4] A. Kirilenko, A. S. Kyle, M. Samadi and T. Tuzun, “The Flash Crash: The Impact of High Frequency Trading on an Electronic Market”, 2011.

[5] P. D. Turney, “Thumbs Up or Thumbs Down?: Semantic Orientation Applied to Unsupervised Classification of Reviews,” *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics (ACL)*, pp.417-424, Philadelphia, July 2002.

[6] J. Wiebe, T. Wilson, R. Bruce, M. Bell, and M. Martin, “Learning Subjective Language,” *Computational Linguistics*, vol. 30, pp.277–308, September 2004.

[7] M. Hu and B. Liu, “Mining and Summarizing Customer Reviews,” *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '04, pp.168-177, Seattle, WA, USA, ACM, 2004.

[8] K. Dave, S. Lawrence and D. M. Pennock, “Mining the peanut gallery: Opinion extraction and semantic classification of product reviews,” *WWW’03: Proceedings of the 12th international conference on World Wide Web*, pp.519-528, New York, NY, USA, ACM, 2003.

[9] B. Pang, L. Lee and S. Vaithyanathan, “Thumbs up? Sentiment classification using machine learning techniques,” *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp.79-86, 2002.

[10] B. Lui, “Sentiment Analysis and Opinion Mining,” Morgan & Claypool Publishers, May 2012.

[11] E. Camria, A. Livingstone and A. Hussain, “The hourglass of emotions,” in *Cognitive Behavioral Systems*, ser. Lecture Notes in Computer Science, A. Esposito, A. M. Esposito, A. Vinciarelli, R. Hoffmann and V. C. Mueller, Eds. Berlin Heidelberg: Springer, vol. 7403, pp.144-157, 2012.

[12] E. Cambria, B. Schuller, Y. Xia and C. Havasi, “New Avenues in Opinion Mining and Sentiment Analysis,” *IEEE Intelligent Systems*, vol.28, no.2, pp.15-21, 2013.

[13] N. Jindal and B. Liu, “Opinion Spam and Analysis,” *Proceedings of the ACM Conference on Web Search and Data Mining (WSDM)*, pp.219-230, 2008.

[14] H.-M. Lu, H. Chen, T.-J. Chen, M.-W. Hung and S.-H. Li, “Financial Text Mining: Supporting Decision Making using Web 2.0 Content,” *IEEE Intelligent Systems*, pp.78-82, 2010.

[15] P. Uhr, A. Klahold, M. Fathi, “Imitation of the Human Ability of Word Association,” *International Journal of Soft Computing and Software Engineering*, vol.3, no.3, pp.248-254, 2013.

[16] A. Klahold, P. Uhr, F. Ansari and M. Fathi, “A Framework to utilize the Human Ability of Word Association for detecting Multi Topic Structures in Text Documents,” *IEEE Intelligent Systems*, November 2013.

[17] C. D. Manning, P. Raghavan and H. Schütze, “Introduction to Information Retrieval,” Cambridge University Press. 2008.

[18] R. Remus, U. Quasthoff and G. Heyer, “SentiWS – A Puplicly Available German-language Resource for Sentiment Analysis,” *Proceedings of the 7th International Language Ressources and Evaluation (LREC'10)*, pp. 1168-1171, 2010.

[19] J. M. Poterba, and L. H. Summers, “Mean Reversion in Stock Prices: Evidence and Implications,” NBER Working Paper No. 2343, July 1989.

[20] B. Qian and K. Rasheed, “Stock Market Prediction with Multiple Classifiers,” *Applied Intelligence*, vol.26, no.1, pp.25-33, 2007.

[21] J. Bollen, H. Mao and X. Zeng, “Twitter mood predicts the stock market”, *Journal of Computational Science*, 2(1), pp.1-8, March 2011.