Quantification of Investor Emotion in Financial News by Analyzing The Stock Price Reaction

Syed Mamoon Ahmed
Department of Computer Science
University of Karachi
University Road, Karachi-75270
Email: phimsyed@gmail.com

Abstract—It is in our nature to measure the intensity of an unseen event by observing its impact on the affected entity. Since reaction of the trader crowd to a financial event in form of stock price movement very closely reflects the severity of that event, words used in the news to describe that event also exhibit similar degree of emotion. Thus, we can assign emotional valence rating to the most relevant words in the news using this reversed relationship of causality instead of assigning assumed emotional valence to words, like it has been the case in many previous works. We have gathered data for financial events of the past from stock exchange and have mathematically analyzed it to quantify the impact of those events. These results are then applied to assign ratings to more than 7000 PoS-tagged English stems extracted from financial news articles for the corresponding events using Natural Language Processing techniques. In this way, a domain-specific word-emotion lexicon has been created.

Keywords: Word Emotion Lexicon, Crowdsourcing, Affective Computing, Sentic Computing.

I. INTRODUCTION

If you were to see a car damaged by an accident on the road, could you tell how bad was the accident? the answer would surely be in the affirmative. If we go by cause-effect terminology; in this instance, Accident is the cause and condition of the wrecked car perceived as effect, we can say that in most cases effect is always directly proportional to the cause[17].

We know from research [6] that stock price of a company reflects the emotion expressed about that company in financial news. Now using our car-accident analogy here, we can say that stock price is the car in this case and financial news is the accident that happens to it. Although, an accident is taken with negative connotations, we are using this only as an example to prove our point about relationship of causality. Of course, if the change in price was positive, we would automatically know that the event is positive too. In this research, we attempt to use this phenomenon of direct proportionality to quantify the investor emotion for English words that occur in the financial news. The deduction flow of our emotional quantification is explained in fig.1 .

Stock price of a company is determined by a fews things. One of them is the activity during the trading day by the investors, the stock price of a company reflects not only how much do the investors think the company is worth but also how much they think it will be worth in the future. It implies that the price of the company stock depends on the perception of the investor. In this way it can be said that stock market is almost like a living organism that reacts to several factors that act upon it [4].

Investor sentiment is the overall attitude of investors toward a particular company or financial market as a whole. It is the feeling or tone of a market as revealed through the activity and price movement of the securities traded in that market. It
is due to this sentiment that investor crowd acts like a swarm, and this swarm is what controls the price of any commodity that is being traded.

If we have the data regarding what events happened in the past and how they were covered by the news, then it can be determined what was the impact of a particular news on the stock price of the company. Fluctuation in the stock price thus introduced is an approximate measure of the human perception regarding the severity of event. Out of many methods that can be used to measure degree of emotional severity of a word. We will take stock exchange as a data source to rate the degree of severity of the language in the news article that caused the change in stock price.

The aspect common in the efforts to assign the emotional rating to Human language words is that they rely on human annotators. However this research attempts to propose a system that is not only scalable and free from annotators biases, but also relies on data that is being generated automatically 24/7 in all parts of the world, is easily accessible and free.

II. BACKGROUND RESEARCH

Working with human emotions in a field of research is as complex as the realm of human emotions itself. However, a few considerable works have been presented in the field of sentiment quantification both in single or multi-dimensional space. As a result of these efforts, we now have a number of word-emotion association lexicons available to use and further build upon. These lexicons may prove themselves to be quite useful in emotion detection algorithms [11] and automatic sentiment analysis [14][5]. Word-Emotion Association Lexicons can be either created manually (i.e. human annotated) or automatically using some statistical technique. For example, probably the very first GPEL (General Purpose Emotion Lexicon) called General Inquirer Lexicon [16] which categorizes words into positive and negative, does so using human-annotation. ANEW (Affective Norms of English Words) is another example of such a GPEL [2]. Since [15] showed the effectiveness of non-expert crowdsourcing systems like AMT (Amazon Mechanical Turk), Other researchers have used AMT to create GPELs [12][11][3].

[1] proposes a method to create DSELs (Domain Specific Emotion Lexicon) to improve the performance of emotion detection related to a specific domain. We use document-level rating in a distributive manner over a BoW (Bag of words) feature model, which is demonstrated in [9] which uses a similar method with IMDB movie dataset.

This research augments stock market psychology with the field of emotional quantification to achieve its goals. Early works in the field of analysis of stock market behavior clearly shows the relation between news and the stock exchange activity [13][10]. These results have led to several research efforts which try to predict the change in stock price of any company that becomes a subject of attention in the latest news on the basis of EMH (Efficient Market Hypothesis) [6][7][8].

III. METHODOLOGY

For the purpose of this research, we have taken publicly listed companies from NASDAQ and NYSE. The choice of focusing on American stock exchanges has been made solely because the stock price data is readily available for longer periods of time and the news coverage for those companies is well-documented and very extensive.

We will start by collecting the data necessary to implement the proposed methodology. In order to extract words from relevant financial news, we first have to detect appropriate events in the history of company stock price suitable for including in our research.

A. Stock Price Data

The gross emotional effect of the event as depicted in fig.2 is the relative change in stock price from point when market starts to take newly available information in account to the point when the price has stabilized[13]. To collect and process the stock price quotation data, we have used Google Finance API. Alternating buy-sell pressure in the stock market creates a very noisy price quotation graph, So even if the stock is moving in a general direction, it becomes very difficult to quantify the effects of a news-inspired movement. A simple solution to tackle this problem is to use SMA (simple Moving Average) of appropriate window. Another benefit of using SMA instead of actual time-series is that it keeps the graph of the derivative from spikes that may have occurred from apparent sudden changes from effect of after-hours trading. For instance fig.2 uses 10-SMA to filter out the noise in real-time price graph.

Figure 2: Initial and final stock price before and after the impact
Following procedure is used to detect appropriate events and then extract the gross emotional impact of those events:

- save the stock ticker (e.g. NASDAQ:TSLA in case of Tesla motor company) of the companies of interest in an array.
- Take each company $c$, retrieve the stock price quotation time-series $TS_c$ and convert the time-series data into 10-SMA i.e. $f(t) = SMA(TS_c)$.
- Derive $f(t)$ with respect to the time $t$ so as to obtain $f'(t) = \frac{d}{dt}SMA(TS_c)$.
- Now that we have the derivative. It is easy to find and locate events of interest. we just have to apply a threshold limit to the $f'(t)$ and mark the time of starting of spike as $t_0$ and time at the ending of spike as $t_f$. both negative and positive change is considered.
- We will now map the $t_0$ and $t_f$ to $SMA(TS_c)$ to find the corresponding and apply (1) to find $\Delta p$. Needless to say, positive $\Delta p$ signifies good news for the company while the negative change signifies lack of trust and fear on part of investor for that particular $(c, t)$.

$$\Delta p = \frac{|p_f - p_0|}{p_0} \times 100 \quad (1)$$

Going through all the companies and events, we now have set $E = \{e_1, e_2, e_3, ... e_m\}$ where, $e$ is an individually identifiable event for company $c$ and time $t$. and $m = c \times t$

### B. Financial News Data

News for time $t$ in the event $e_i(c, t)$ from multiple authentic and highly cited sources is scraped from the web covering the same event that has caused the impact $\Delta p$ on price of company $c$. Python library ‘Goose’ has been used for this purpose. Before $\Delta p$ can be applied to the individual stems, we must pre-process the data.

- Article titles and bodies are stripped of any symbols, stop words, Numerical values or Proper Nouns.
- PoS tagging is done using maximum entropy algorithm from Python NLTK library.
- After PoS tagging of all non-stop words the words are stemmed using snowball algorithm to reduce the dimensionality.

As a result of the above process, We now will have obtained set $A_i(c, t) = \{s_1, s_2, s_3, ... s_n\}$ where $A$ is the set of all stems extracted from news related to event $e(c, t)$ on which we can apply $\Delta p$. and $n$ is the number of stems extracted from the articles pertaining to corresponding event.

### C. Finding Cumulative Emotional Valence

Complete algorithm employed to apply the $\Delta p$ to the bag of stems and saving the final result to $S$ and $P$ is shown in fig.3. Where $S = \{s_1, s_2, s_3, ... s_q\}$ is the set of all unique stems in the dataset $E$ and $P = \{\Delta p_1, \Delta p_2, \Delta p_3, ... \Delta p_q\}$ The resulting set $R$ is a key-value dictionary which can be represented as in (2).

$$R : S \to PU\{\epsilon\} \quad (2)$$
The exact procedure to apply $\delta p$ to each stem in article set $A(i)$ for the event $e(c, t)$ is shown in fig.4. As for final result $\Delta p_1$ for $s_1$ for example is the sum of all $x \times \delta p$ for every event $e(c, t)$ that the stem $s_1$ occurs in. Here $x$ is the number of times $s_1$ has occurred in event $e$.

IV. RESULTS

In this section we will see and discuss the results obtained using the method described in the last section. In fig.5 you can see the emotional rating distribution ($\Delta p$) upon the stems occurring in all the documents analyzed. More than 7000 stems have been rated, and results have been grouped by basic parts of speech which hold the most emotional temperature in any document i.e. Adjectives, Adverbs, Verbs and Nouns. The graph has been color coded, with Noun graph in Green, Verb graph in Blue, Adverb graph in Red and Adjective’s graph in Cyan. The final rating for each stem can be summarized in the following formula:

$$\Delta P_s = \sum_{i=0}^{q} \delta p_i \times y$$ (3)

In above equation, $\Delta P_s$ is the cumulative rating for stem at the end of the iteration of all the events, while $q = n \times n \times y$ is the number of times stem $s$ occurs each event. if the stem does not occur in any article of a particular event $(c, t)$ then its value of $\delta p$ is considered '0'. Now coming back to the graph, We can deduce a few things by critically analyzing the resulting graph in fig.5

- It proves the hypothesis that some words occur in the negative news more than positive news and vice versa.
- The differential ranking method between two set of articles which are opposite in emotions works.
- The frequency distribution of the differential emotional score complies with the Zipf’s Law.
- Like the frequency of words, emotional frequency of words is also concentrated in the far ends of the word spectrum, proving how humans like to express a wide range of emotions with a combination of as fewer words as possible.

Even though it can be seen that the method works, it can also be seen that it is domain specific (i.e. specific for financial news) and dimension specific (i.e. it can only measure greed and fear). Also it does not take into account any insider trading or the surprise factor of the news (i.e. more surprised the investor, higher the impact).

Examples of words that are on the positive side include: high, well, better, grew, jump, short, best, invest, ahead etc., while some of the words on other end of spectrum are: plunge, drop, charges, crisis, complaint, worst, disappoint, prosecutor, explode, wipe, dispute, downgrade, worst, sour, scandal, lowest, unfortunate etc. since numerical values were not excluded from the ratings, we can also see “2017” and “2016” as high-scorers, which can be explained away by the fact that most of the news in the corpus is positive and is based on many instances of successful quarterly financial reports of companies.

While many ratings can be explained by common...
sense, others cannot. For example, numerical values ‘1’ and ‘3’ score very high on the spectrum; and words like ‘and’, ‘also’, ‘last’ also show very positive rating. which may be due to the unbalanced data, or pure chance. This may be cleared upon further investigation of the matter.

V. CONCLUSION AND FUTURE WORK

This work has been successful in providing some evidence in favor of the initial hypothesis and completing the objective. Also due to its mathematical simplicity, it is easy to understand and implement. The weakness of this work is its rudimentary nature, and its domain specificity(which may also be counted as a strength). We have discovered a new and generally implementable crowdsourcing model but have only proved it for one domain and dimension. It goes without saying that as more data is collected from a more diverse set of sources and for a more diverse set of stocks, the ratings will improve in generalization and accuracy.

In order to further research in this direction, other kind of databases can be analyzed for such ratings so that the words related to that realm can also be ranked. We can also implement a better method to analyze the stock trends and to apply the rating in a uniform fashion as opposed to the very fundamental way it has been done in this research.

REFERENCES

[1] A. Bandhakavi, N. Wiratunga, D. Padmanabhan, and S. Massie. “Lexicon based feature extraction for emotion text classification”. In: Pattern Recognit. Lett. 93 (2017), pp. 133–142.

[2] M. M. Bradley and P. P. J. Lang. “Affective Norms for English Words ( ANEW ): Instruction Manual and Affective Ratings”. In: Psychology Technical.C-1 (1999), p. 0. arXiv: arXiv:1011.1669v3.

[3] E. Estellés-arolas and F. González-ladrón-de-guevara. “Towards an integrated crowdsourcing definition”. In: J. Inf. Sci. 38. March (2012), pp. 189–200.

[4] E. Fama. “The Behaviour of Stock-Market Prices”. In: J. Bus. 38.1 (1965), pp. 34–105.

[5] X. Fang and J. Zhan. “Sentiment analysis using product review data”. In: J. Big Data 2.1 (2015), p. 5.

[6] G. P. C. Fung, J. X. Yu, W. Lam, G. P. C. Fung, J. X. Yu, and W. Lam. “News sensitive stock trend prediction”. In: Adv. Knowl. Discov. Data Min. 2336 (2002), pp. 481–493.

[7] G. Fung, J. Yu, and H. Lu. “The Predicting Power of Textual Information on Financial Markets.” In: IEEE Intell. Informatics Bull. 5.1 (2005), pp. 1–10.

[8] A. Khadjeh Nassirtoussi, S. Aghabozorgi, T. Ying Wah, and D. C. L. Ngo. “Text mining for market prediction: A systematic review”. In: Expert Syst. Appl. 41.16 (2014), pp. 7653–7670.

[9] M.-C. de Marneffe, C. D. Manning, and C. Potts. “Was it good? It was provocative.” Learning the meaning of scalar adjectives”. In: ACL ’10 Proc. 48th Annu. Meet. Assoc. Comput. Linguist. July (2010), pp. 167–176.

[10] M. L. Mitchell and J. H. Mulherin. “The Impact of Public Information on the Stock Market”. In: J. Finance 49.3 (1994), pp. 923–950.

[11] S. M. Mohammad and P. D. Turney. “Crowdsourcing a Word-Emotion Association Lexicon”. In: Comput. Intell. 29.3 (2013), pp. 436–465.

[12] S. M. Mohammad and P. D. Turney. “Emotions evoked by common words and phrases: using mechanical turk to create an emotion lexicon”. In: CAAGET ’10 Proc. NAACL HLT 2010 Work. Comput. Approaches to Anal. Gener. Emot. Text June (2010), pp. 26–34.

[13] D. K. Pearce and V. V. Roley. “Stock Prices and Economic News”. In: NBER Work. Pap. Ser. 1296 (1984).

[14] C. Quan and F. Ren. “An exploration of features for recognizing word emotion”. In: Proc. 23rd Int. Conf. Comput. Linguist. August (2010), pp. 922–930.

[15] R. Snow, B. O. Connor, D. Jurafsky, A. Y. Ng, D. Labs, and C. St. “Cheap and fast - but is it good? Evaluation non-expert annotations for natural language tasks”. In: Proc. Conf. Empir. Methods Nat. Lang. Process. October (2008), pp. 254–263.

[16] P. J. Stone, D. C. Dunphy, M. S. Smith, Ogilvie, and D. M. “The General Inquirer: A Computer Approach to Content Analysis”. In: American Journal of Sociology 73.5 (1966), pp. 634–635. eprint: https://doi.org/10.1086/224539.

[17] L. C. Yu, J. L. Wu, P. C. Chang, and H. S. Chu. “Using a contextual entropy model to expand emotion words and their intensity for the sentiment classification of stock market news”. In: Knowledge-Based Syst. 41 (2013), pp. 89–97.