A Study of the Relationship Between Company Stock Price and Twitter Sentiment of Product Releasing Event

Hao Gao*, An Shi*
Department of Computer Science, Illinois Institute of Technology
Chicago, IL
hhgao16@hawk.iit.edu ashi@hawk.iit.edu
* Authors made equal contribution
All our work is open source at https://github.com/anshipku/Events_vs_Stock.

Abstract: People are getting more and more interested in predicting stock price using Twitter information. However, few of them focused on particular events, at which period the stock price may have big fluctuations. In this paper, we carried a study of the relationship between stock price, specifically daily residual and Twitter sentiment, analyzed by *SentiWordNet 3.0*, when a new product is released. Our results showed a similar trend of stock price and Twitter sentiment around the product releasing date.

1. Introduction

Every big event (new products releasing, financial statement report) will have effects on the company's stock price somehow, since the stock is a very sensitive information indicator. Since certain product only could have one true value, there always will be mispricing towards the product, so there are some chances to make money out of these mispricing. Before these big events, people will have different reaction towards the events. Sometimes, people overreact toward the new product, sometimes they underestimate. When people overact to the news, the company's stock tends to go up high than its true value and vice versa. This means there is a chance to mispricing the product. For example, since iPhone 4 was released, which is an outstanding product, people would have high expectation towards iPhone 4s. After iPhone 4s was released, though it is a good product, it failed to fulfill the anticipation of "change the world again". Thus, before the product releasing, the price of the stock will go uphill and after the event, the price will steadily drop back, till its true price. In this paper we want to figure out whether there is a correlation between people's reaction towards these new products and stock price.

2. Related work

We started with the famous paper *Twitter mood predicts the stock market*[1], which is almost the first paper we read from computer science point of view, using machine learning technique to analysis finance event. However, there are some apparent flow with this paper, for example, they only picked a particular period time to do the test. But it inspired the idea we used in our experiments. There is another paper named *Stock Prediction Using Twitter Sentiment Analysis*[2]. They have almost adopted the whole experiments process of the former paper and did some mineral change with it, for instance, they used different sentiment dimensions to predict the stock price.

Correlation Financial Time Series with Micro Blogging Activity[3] is different from the above. They first tried to give a more detailed way of analyzing Twitter data. They classified the different actions adopted by Twitter users, such as number of re-tweets, number of tweets with geo-location, etc. Then, depending on this, they wanted to figure out the correlation between Twitter and stock price (and some other financial ratio, like earning and price).

And, of course, the behavior finance area beginning way ahead of Twitter and even machine learning area. There are two major hypothesis to describe individual's sentiment affect their market behavior. Under efficient market hypothesis, it is the market who determines the price of certain product. But under Hiller's hypothesis, people's reaction may be randomly distributed and guide their market behavior. There are many related papers on this area, however, due to limited time, we couldn't go through all of them. *Investor Sentiment in the Market*[4] by Malcolm Baker, give us some basic information about what kind of models are generally adopted in measuring people's feelings and stock price in finance area.

We want to narrow down the research target from a general window, people's sentiment and stock price, down to a particular relationship after an event. Thus, the results are expected to be filtering out a lot of noise of vary kinds of information.

3. Data

3.1 Stock Data

There are two major parts of the data we used in this paper. Frist of all, we collected appropriate events that could fit in our purpose and model. The model we used is called *Capital Pricing Asset Model*(CAPM), since we are want to calculate daily expect return of single stock. CAPM assumes that stock inside certain asset allocation has a correlation or beta with the bench mark. The bench mark is calculated out through minimizing the variance and covariance of target assets. We only picked US IT company that listed in the Nasdaq 100, since we used Nasdaq 100 as the benchmark or the base line and because of the nature of CAPM model. The data we used to calculate the expect return and market beta were all derived from *Yahoo Finance*. The events we picked are listed in **Table 1**.

Company	Product	Release Date	Number of Tweets
Apple	iPad 1	04/05/2010	96503
	iPad 2	03/11/2011	9980
	iPhone 4s	10/14/2011	15018
	iPad 3	03/16/2012	27028
	iPhone 5	09/21/2012	25712
	iPad 4/mini	11/02/2012	0
Amazon	Kindle Fire 1	11/15/2011	1289
	Kindle Fire 2/HD	09/14/2012	853
Google	Nexus 7	07/15/2012	1850
	Nexus 4/10	11/13/2012	0

Table 1. The product list of our research, the number of tweets shows the size of the data set.

We mainly focused on observing the stock price and sentiments changing around the product releasing day. The time range window we are observing are typically from 10 days before the releasing date and 10 days after it.

3.2 Twitter Data

We got all our Twitter data from Professor Aron Culotta of IIT. They are collected using Twitter Streaming API and contained approximately 1% sample of all global tweets from 2009 to 2012. We first removed all the tweets that are not tweeted by users in US, by looking at the user's time zone and keeping only those in the US. We then filtered the tweets by keyword matching in the above column(like "iPad") within the time interval of a specific product we are interested, which is typically from 10 days before the product release to 10 days after the release. For a specific time interval, we only searched the product name without the model number, since we assumed that even people were tweeting about other models of the same product, it still had an effect on the coming new product. Furthermore, people always don't know the exact model number before a product is actually released. We also removed tweets containing strings matching a specific product, but not about the product, for example, "helipad" has "iPad", but it is apparently not about iPad.

The number of tweets after filtering are listed in **Table 1**. Noticing that there are some days for some product that has no tweet at all, which is due to the lack of data. The two most complete ones are iPad 1 and iPad 3, we had to narrow some of experiments to these two products in order to get conclusive results.

4. Methods

4.1 Stock Data

First we need to deal with the stock price. Market is an information chaos and we want to try our best to filter out various kinds of noise. The finance method we adopted is event study. We set a target event that occured in the past, like new product releasing of a company. Then set the event day as day 0 and we are going to examine the response of the stock from 10 days before to 10 days after the event (-10 to +10), we might estimate the characteristic line in the period from 70 days before to 11 days before(-70 to -11). The chart of this is called characteristic line. The characteristic line gives us an expectation of what the stock's return should be on a particular day, given that the market has produced a particular rate of return.

We used CAPM model to calculate expect return, since we needed the daily based return. We first calculated the market beta of certain stocks based on the past 3 months (70 days) historical stock return and benchmark return. Using linear regression:

$$R_i = \alpha + \beta R_B + T \qquad (1)$$

R_i And R_B here are the historical return of certain stock and the benchmark. The stock, of course, was chosen based on events that we picked. The benchmark here we used is Nasdaq 100, since the events we picked up were generally technical companies' product releasing and Nasdaq 100 is generally the benchmark of technology companies. Based on the historical return, we used linear regression to figure out the market beta and intercept. The T here is the T-bill rate and under the current market situation, it is too small to be noticed. Then we can calculate the expect return of the stock based on

$$ER_i = \hat{\alpha} + \hat{\beta} R_B \quad (2)$$

The $\hat{\alpha}$ and $\hat{\beta}$ are the result of the previous regression. Noticing that this time we directly ignored the T-bill.

Since we are interested in the response of the stock to the event, we would like to determine if the stock's return in the vicinity of the event are above or below what we would expect to see in the light of the performance of the market. Thus, our measure of response $\varepsilon_{i,t}$, on a given day is the difference between the stock's actual rate of return and its conditional rate of return.

$$\varepsilon_{i,t} = r_{i,t} - E\left(r_i | r_{n,t}\right) \quad (3)$$

To measure the stock's overall reaction to the event, we might accumulate the responses going from day -10 to day +10

$$\sum_{i=-10}^{+10} \varepsilon_{i,t} = E_{i,+10} \quad (4)$$

Individual stock is subject to the flow of a variety of types of information, some of which will push the stock up and some of which will push the stock down. Thus, the response to the particular event of interest may be masked by the responses to the other prices of information. To handle this problem, we collect a large sample of stocks that have in common the incidence of a particular event, which in this case, is the product releasing. The releasing may come at different calendar times for the different stocks, but for each stock we compute the accumulated response relative to day 0, the day the releasing is averaged over all M stocks in the sample as follows to obtain what is called the cumulative average excess return:

$$\bar{E}_t = -\frac{\sum_i^M E_i,t}{\mu} \quad (5)$$

Since the only thing the stocks in the sample have in common is the event, the other factors that are influencing their prices should cancel out in the averaging. The movement in $\bar{E}_t$ as we approach and accuracy of the response of stock prices to the particular event of interest.

4.2 Twitter Data

4.2.1 SentiWordNet 3.0

SentiWordNet3.0 is a lexical resource for opinion mining[5], it assigns to each word in the list, under different contexts and parts of speech, three sentiment scores: positive, negative, neutral, each polarity ranges from 0 to 1.0, and all three adds up to 1.0. However, it is time consuming and difficult to analyze the context of each word in a tweet, so we adopted a similar approach as *Mohamed Rohaim*[6], by averaging the sentiment score of each polarity for a specific word with a part of speech. For example, the word "good" has 21 meanings when used as "a"(adjective and adverb), 4 meanings as "n"(noun), the averaging results are shown in **Table 2**. We then removed the neutral words whose positive and negative score all equal to 0 since they will not contribute the sentiment analysis. This gives a total vocabulary of 41593 words.

Word#POS	Positive Score	Negative Score	Number of Occurrence
good#a	0.6190	0.0060	21
good#n	0.5000	0.0000	4
bad#a	0.0179	0.0661	14

Table 2. An example of the dictionary of *SentiWordNet 3.0*.

4.2.2 Sentiment Analysis of Tweets

After removing punctuations and URLs in tweets, we tagged each word its part of speech by *Stanford POS Tagger* and converted the detailed labels to three types: "n" for noun and pronoun, "v" for verb and "a" for any type left, to match the types used in *SentiWordNet 3.0*. We then matched the words on each tweet with the vocabulary, and calculate its polarity score by equation (6).

$$A\ Tweet's\ Polarity\ Score = \frac{\Sigma\ Matched\ Word's\ Polarity\ Scores}{\#\ of\ Mathched\ Words\ in\ that\ Tweet} \quad (6)$$

Thus, each tweet has three sentiment scores: positive, negative, neutral. During a particular day, we add up all the related tweets' sentiment scores and normalize it by dividing the number of tweets. We ignored those tweets who does not contain any of the words in *SentiWordNet 3.0*.

$$A\ Day's\ Polarity\ Score\ by\ Tweets = \frac{\Sigma\ Tweets'\ Polarity\ Scores}{\#\ of\ Tweets\ in\ That\ Day} \quad (7)$$

$$A\ Day's\ Polarity\ Score\ by\ Words = \frac{\Sigma\ Matched\ Words's\ Polarity\ Scores}{\#\ of\ Mathched\ Words\ in\ That\ Day} \quad (8)$$

A specific day's sentiment is presented by the difference between positive score and negative score. Moving averages of 3 and 5 days are also calculated to reduce the noise by the following formula:

$$A\ Moving\ Average\ of\ n\ Days = \frac{\Sigma\, S_M + S_{M-1} + \cdots + S_{M-(n-1)}}{n} \quad (9)$$

We also labeled each tweet to be either positive if its positive score is greater than negative score, or negative if its negative score is greater than positive score. There's almost no neutral tweets as the positive and negative scores are hardly the same. Then we calculated the percentage of positive tweets in each day.

$$Percentage\ of\ Positive\ Tweets = \frac{\#\ of Positive\ Tweets}{\#\ of Positive\ Tweets + \ \#\ of Negative\ Tweets} \quad (10)$$

So, for each day, we want find out whether there is a linear relationship between the excessive return of the company's stock and people's sentiment.

5. Experiments

5.1 Stock

Based on the event we located, we found out the stock price back and forth 10 days from day 0. Since the market close at weekend, there is no stock prices during those days. We used Monday and Friday's stock price as continuous data. Since we calculated the moving average of Twitter sentiment, the effect of weekends has already be considered in the following weekdays.

As it be mentioned before, we used the stock's price and benchmark's price for past 3 months to calculate the market beta of certain stock. The use the market beta and the benchmark's price around the day 0 to calculate the difference between the expect return and the real return of the stock.

Here is an example in **Table 3**. March 11 is the event day, which the day that Apple release iPad 2. The stock price and bench mark price we used is the daily adjust return. Market beta is calculate out by the linear regression:

	Intercept	**Adj Close**
Coefficients	-431.465	0.354597
Standard Error	30.01336	0.012394
t Stat	-14.3758	28.6104
P-value	4.27E-21	1.14E-36
Lower 95%	-491.5	0.329805
Upper 95%	-371.429	0.379388

Table 3. An example of calculation of daily adjust return.

Adj Close here is the slope resulted from the linear regression, which is the market beta we need.

Daily residual here is the difference between the stock's real return and the expect return. Based on the equation:

$$\varepsilon_{i,t} = r_{i,t} - E\left(r_i | r_{B,t}\right) \quad (11)$$

The $r_{B,t}$ here is the real return from the benchmark, which corresponds to QQQ event day adjust return. The results of iPad 3 are shown in **Table 4** as an example.

Date	**APPL Event Day Adj Close**	**QQQ Event Day Adj Close**	**Daily Residual**
02/25/2011	338.60	2346.29	1.060
02/28/2011	343.52	2350.99	5.096
03/01/2011	339.72	2315.26	8.018
03/02/2011	342.46	2326.77	8.593
03/03/2011	349.69	2371.76	7.359
03/04/2011	350.12	2359.96	10.009
03/07/2011	345.61	2328.07	11.498
03/08/2011	346.00	2337.55	10.105
03/09/2011	342.80	2322.69	9.700
03/10/2011	337.15	2284.29	11.274
03/11/2011	342.33	2299.26	13.638
03/14/2011	343.86	2290.72	16.774
03/15/2011	335.95	2259.62	14.715
03/16/2011	320.95	2202.97	10.372
03/17/2011	325.46	2225.24	10.693
03/18/2011	321.59	2221.07	7.607
03/21/2011	329.99	2262.70	8.176
03/22/2011	331.84	2257.96	10.917
03/23/2011	329.88	2270.50	6.598
03/24/2011	335.50	2312.09	4.394
03/25/2011	341.89	2316.36	9.981

Table 4. The results of daily residual of iPad 3.

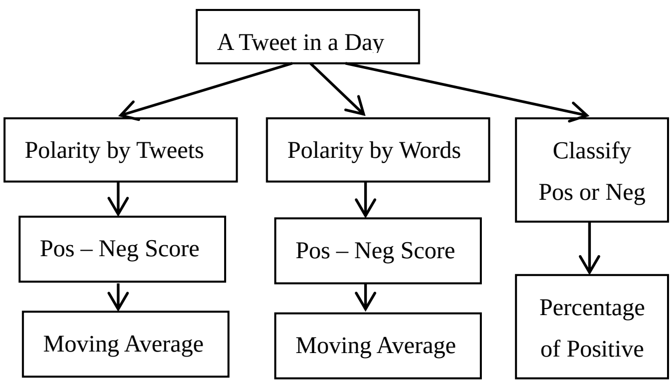

Figure 1. Workflow of Twitter data analysis.

5.2 Twitter

We first calculated number of tweets each as well as the percentage of on optic tweets. We then calculated each day's polarity by tweets, words and percentage of positive tweets in that day. **Figure 1** shows the workflow of Twitter data analysis.

5.3 Correlation between Stock and Twitter

By analyzing the stock daily residual and moving average of Twitter sentiment score, we expected to see some interesting results.

5.4 Linear Regression between Stock and Twitter

To explore the relationship between the stock price and Twitter sentiment, we ran a linear regression of Daily Residual with positive and negative score of each day.

6. Results

6.1 Average Daily Excessive Return

The daily average residuals in **Figure 2** are the general differences between the stock prices and the benchmark. By adding up different stocks' daily excessive return, we hope to cancel out the noises bury inside different stock prices, since the price changing of a certain stock probably due to the different events. The ideal object is to cancel out all the noise and make the stocks only response to one particular event. In this case, the event is new product releasing.

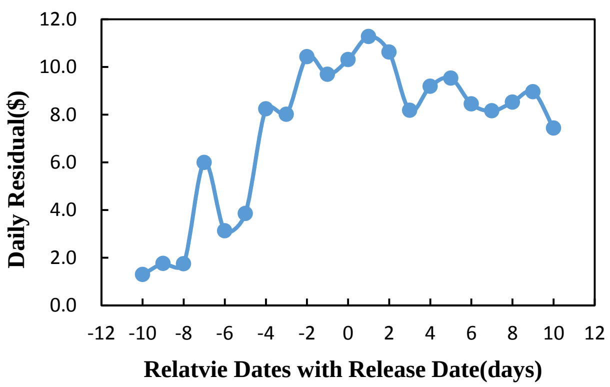

Figure 2. Daily average residual of all products.

6.2 Number of Tweets

Figure 3 clearly shows that the number of tweets reached a peak at the releasing date, followed by two more peaks during 5 days after release. The percentage of iPad 1 tweets over the total number of tweets in the same day clearly showed a peak at the releasing date. So we can see people are tweeting more about the product during releasing.

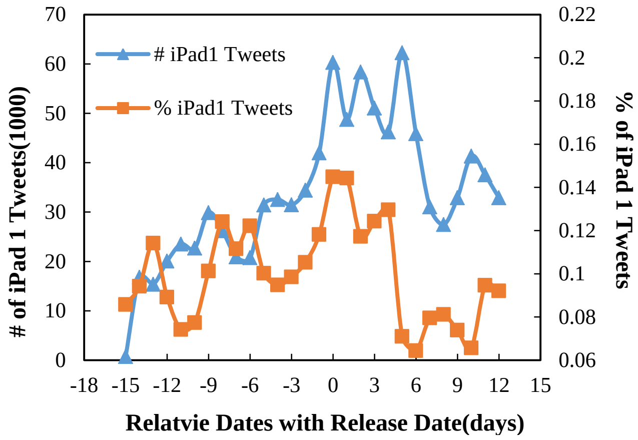

Figure 3. Number of iPad 1 tweets and the percentage of iPad 1 tweets over dates.

6.3 Different Methods of Handling Twitter Data

The different method of handing Twitter data as indicated in **Figure 1** gave us similar results. And moving average of 3 and 5 days also shows similar trend of data, however, the more days one take into the average, the more smooth the curve is. To simplify this paper, we will only use polarity by total tweets and a 3 day moving average.

6.4 Correlation between Stock and Twitter

After analyzing the correlation and linear regression between stock and Twitter. We got a general conclusion: Making money is not easy. There is no clear linear relationship between the stock price and the sentiment in our data. Though there is some interesting similar pattern between the stock price and sentiment scores, but, clearly, it needs more analysis.

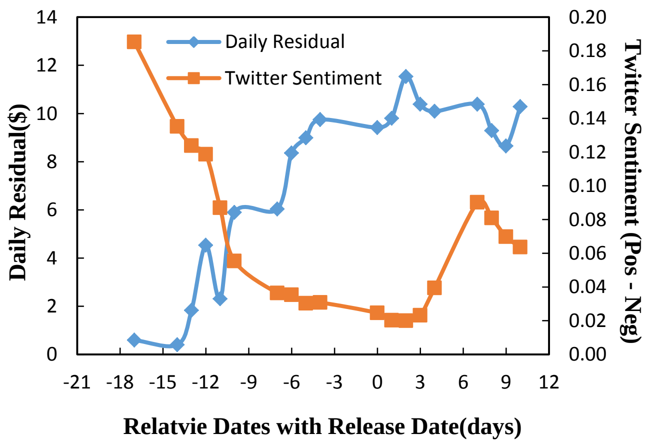

Figure 4. Daily residual and Twitter Sentiment of iPad 1. The Twitter sentiment used here is the difference between positive and negative score of each day. To minimize noise, the sentiment difference presented by a moving average of three days.

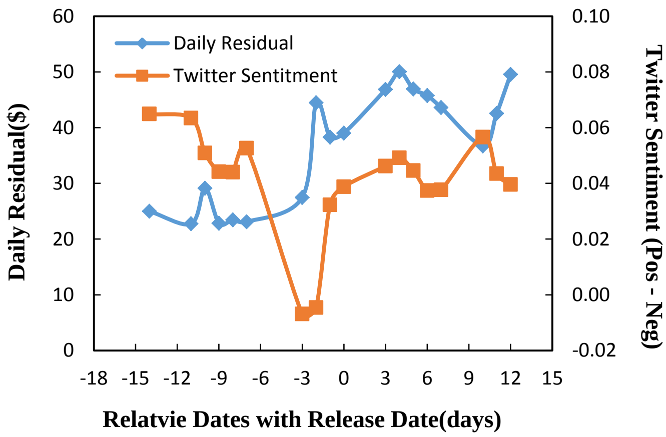

Figure 5. Daily residual and Twitter Sentiment of iPad 3. All data are similar to Figure 4.

However, there is an extremely interesting pattern that worth to be noticed. **Figure 4** and **5** shows the daily residual and Twitter sentiment of iPad 1, iPad 3.

From these figures we can see:

1. The daily residual almost perfectly matches the market efficient theory. Taking product releasing as an gradually information releasing process, before each event, market didn't know what was the quality of this product and what should be the true price of the company's stock. After the product releasing, the market gathered the missing information and relocated the company's stock price.

2. One also can see the sentiment changing in Twitter. Almost before every product releasing, people's sentiment went straight down to the bottom, and gradually bounced back after the product was released.

3. Almost each inflection points, they are corresponding to each other.

The results are not against both hypothesis. The stock price pattern is almost classic stock price changing patter, under semi-strong market. Also, under Hiller's theory, people's reaction towards certain event should be reversed to the stock price, which is also reflected on our results. And, between a couple more days before and after the event, the sentiment changing and the price changing are in the same trend. However, as getting close to the event day, it displays the reverse pattern. Then after a few day, they back to the same trend. Most of these patterns' inflection points are identical.

The results indicate that under this particular condition, the two hypothesizes seem to have some correlation with each other and need to be studied in the future. Even though there is no clear linear relationship between each other, one cannot denied there is a pattern.

6.5 Linear Regression between Stock and Twitter

The linear regression results are show in **Table 5**. The low regression score (Coefficient of Determination) shows that there's no clear linear relationship between those data, as also indicated by **Figure 4** and **5**.

R^2	Total Tweets	Total Words
iPad 1	0.37	0.42
iPad 3	0.00	0.18

Table 5. Linear Regression of stock and Twitter sentiment. The total tweets and total words are two different ways of handling sentiment.

7. Future Work

First of all, there are already a lot of systematically investor sentiment study in finance area. We need to combine their results with the methods we adopted here. For example, in this paper, we used the *SentiNetWord 3.0* list to calculate the sentiment of Twitter. However, we need to know, in general, what kind of methods does finance people measure the investor's sentiment and whether they have overlapped area with *SentiNetword 3.0*. We need also to consult with psychology expert to see the standard measurement they adopted to gauge investors. Then we could decide what kind of methods we should adopt to do the future work.

Secondly, the machine learning algorithm should be updated to another level. For example, if a company released a welcomed product in the past, say half a year, customer's sentiment towards this company should be generally positive. Thus, they anticipated the next product to be good. All these information would be reflected on the stock price in the new products releasing day. On the other hand, if the company released a bad product in the past, people's expectation should be low in general and all of these would be reflected in the stock price. Thus, it is clear involved with a detailed classification problem besides the words of sentiment classification. These problem could be solved by the higher level of machine learning algorithm, however, the results' accuracy will depend on the parameter we give.

Finally, we need to dig more into Twitter. Different people have different influence on others. Certain tweets could be re-tweeted for countless times. Certain blogs could be mentioned a lot of times. Clearly, these tweets and articles have strong impact on others. Thus, how do we place different weight on these events? What kind of social media study could we adopt in our research? All these questions are believed to help us to get closer the true relationship between stock price and people's sentiment around certain event.

8. Authors' Contributions

The two authors generally work together on the ideas, product search and results analysis. Each of them has their own focus during the data analysis and document composing. **Table 6** lists the detailed contributions. In general they have equal contributions to this work.

Job List	Hao Gao	An Shi
Ideas	√	
Product Search	√	O
Stock Data Acquisition	√	
Twitter Data Acquisition	O	√
Stock Data Analysis	√	
Twitter Data Analysis		√
Correlation	√	√
Regression		√
Project Update Presentation	O	√
Project Final Presentation	√	O
Final Report	√	√
All about Gibhub		√

Table 6. List of author's contributions. "√" indicates major contribution, "O" indicates minor contribution.

9. Acknowledgement

We thank Professor Aron Culotta sincerely for introducing us into the machine learning world and giving us this opportunity to work together on this interesting project, as well as discussing ideas with us and providing all Twitter data. Inspired by this project, we would like to do more work in this field in the future.

10. References

[1] Johan Bollen1, Huina Mao, Xiao-Jun Zeng. Twitter Mood Predicts the Stock Market. DOI: 10.1016/j.jocs.2010.12.007.

[2] Anshul Mittal, Arpit Goel. Stock Prediction Using Twitter Sentiment Analysis.

[3] Eduardo J. Ruiz, Vagelis Hristidis, Carlos Castillo, Aristides Gionis. Correlating Financial Time Series with Micro-Blogging Activity. *Proceedings of the fifth ACM international conference on Web search and data mining*, Pages 513-522.

[4] Malcolm Baker, Jeffrey Wurgler. Investor Sentiment in the Stock Market. *Journal of Economic Perspectives*, Volume 21, Number 2, Spring 2007, Pages 129–151.

[5] Andrea Esuli, Fabrizio Sebastiani. SentiWordNet: A Publicly Available Lexical Resource for Opinion Mining. *Proceedings of the 5th Conference on Language Resources and Evaluation*.

[6] Alaa Hamouda, Mohamed Rohaim. Reviews Classification Using SentiWordNet Lexicon. The Online *Journal on Computer Science and Information Technology (OJCSIT)*, Volume 2, Number 1.)