Sentiment Analysis of Investor Opinions on Twitter

by

Brian Dickinson

Submitted in partial fulfillment of the requirements for Major Honors in Department

Houghton College, Houghton, New York

May, 2015

Honors Committee

Dr. Wei Hu: Signature:____________________________

Dr. John Rowley: Signature:____________________________

Prof. Kenneth Bates: Signature:____________________________

Table of Contents

Abstract

The rapid growth of social networks has produced an unprecedented amount of user-generated data, which provides an excellent opportunity for text mining. Sentiment analysis, an important part of text mining, attempts to learn about the opinion of the author of a text through its content. Such information is particularly valuable for determining the overall opinion of a large number of people. Examples of the usefulness of this are predicting box office sales or stock prices. One of the most accessible sources of user-generated data is Twitter, which makes the majority of its user data freely available through its data access API. In this study I seek to predict a sentiment value for stock related tweets on Twitter, and demonstrate a correlation between this sentiment and the movement of a company's stock price in a real time streaming environment. Both n-gram and "word2vec" textual representation techniques were used alongside a random forest classification algorithm to predict the sentiment of tweets. These values were then evaluated for correlation between stock prices and Twitter sentiment for that each company. There were significant correlations between price and sentiment for individual companies, these correlations were not however consistent between companies. Some companies such as Microsoft and Walmart showed strong positive correlation, while others such as Goldman Sachs and Cisco Systems showed strong negative correlation. This suggests that consumer facing companies are affected differently than other companies. Overall this appears to be a promising field for future research.

1. Introduction

Over the last several years there has been an explosion of growth and new activity in social networking. Various companies such as Facebook, LinkedIn, Reddit, Pintrest, and Twitter have grown exponentially in recent years. The amount of data exchanged between users on these sites is staggering. On Facebook alone on an average day in 2014 there were 4.75 billion items shared, 4.5 billion items "liked", and 300 million photographs uploaded. That translates to over 500 terabytes of data generated by Facebook users on a single day. [1] There is an incredible amount of useful information about individual opinions, feelings, and relationships contained in these transactions, but the loosely structured nature of human communication makes harnessing this data a challenge. In order to make sense of the large portion of this data which is text based, Natural Language Processing tools can be used to rigorously categorize user generated text. One of these tools for determining useful information from massive data sources such as Twitter is sentiment analysis.

Sentiment analysis focuses on determining the opinion of a speaker on the particular topic about which he is speaking. The most basic structure for sentiment analysis is a single word, unfortunately based on sentence structure and words with context dependent meanings, techniques that ignore sentence structure or bag of words models often fail on smaller texts. A solution to this is constructing parse trees which identify the structure of a sentence as a binary tree by separating distinct phrases. In this case using the sentiment of each word in the tree can take into account clause structures and the possibility of multiple meanings. In cases where a larger text must by analyzed, it can be treated as a collection of smaller phrases, or as a larger bag of words. Opinions are usually classified somewhere between positive or negative often with some stratification between the two. This can be done numerically or categorically. When division is categorical it usually distinguishes between positive, negative, and sometimes neutral sentiments, otherwise the numerical classification falls somewhere on a continuum between positive and negative. These classifications can be used to determine and aggregate the sentiment of a large number of authors on a given topic.

Because sarcasm, and even simple negations can completely reverse the predicted sentiment from the actual opinion represented, parse trees are the most accurate method of determining the sentiment of sentences. [2] One of the most recent methods of sentiment analysis, published by Stanford, has advanced features which allow it to recognize some of the most difficult features of human languages. It can identify subordinating conjunctions such as “if”, “after”, “once”, “since”, and “because” and weight towards the portion of the sentence after these since it is more important to the meaning of the sentence. Similarly it can identify the purposes of certain modifiers such as “very” and “incredibly” and use them to weight the sentiment of the words they modify rather than giving them a sentiment of their own. Finally it can handle negations quite well, and is usually accurate in switching the sentiment when something is negated [3]. All of these features make this new system for sentiment analysis very promising as a foundation for applications based on sentiment.

This Stanford model makes use of deep neural networks. Deep neural networks are an expansion on early neural networks such as Perceptrons. Advances in hardware processing speeds, particularly graphics processing units, as well as an increasing interest in parallel processing have brought a resurgence in the use of artificial neural networks by enabling the addition of hidden layers of neurons, and backpropagation. The additional layers allow these models to become more highly non-linear fitting closer to the data, while backpropagation enhances training efficiency on labeled data in deep neural networks. Such networks have won numerous competitions in pattern and image recognition competitions over the last five years, and appear to have great promise in improving classification accuracies in most areas of machine learning.

One topic of user sentiment that can be easily checked for correlations between public opinion and public behavior is that of stock price prediction. There are two basic methods for predicting whether the price of a given stock will rise or fall, fundamental, and technical analysis. Fundamental analysis relies on the financial data of the company to make assessments of financial stability, growth potential, and inherent value. This value can be matched against the current market price. If the estimated real value is higher than the current market price, it is believed that the company is undervalued and that the stock is more likely to rise than to fall. Similarly if the estimated value is lower than the market price, it is assumed that the stock is overvalued and that it is more likely to fall than to rise. [4]

Technical analysis takes a different route. Rather than focusing on financial data, technical analysis uses historical price data to make predictions about the expected direction of price change in the future. Frequently observed patterns that appear to occurring such as head and shoulders or double tops, as well as recent trends such as channels and uptrends are used to predict future prices. Another way technical analysis seeks to predict prices is through observing the behavior of others. One factor in technical analysis is buying and selling at the same time as company insiders, and buying and selling opposite odd-lot traders. The idea behind this is that company insiders are best acquainted with the company's prospects, and make the most educated buying and selling decisions. Meanwhile odd-lot traders are almost always individual rather than professional traders and generally lack a strong investing background. Their trades are negatively correlated with trades by company insiders, and generally result in buying when prices are high, and selling when prices are low. By buying and selling opposite them, one can often buy when prices are low and sell when prices are high. [5] None of these methods guarantee success, but they are generally accepted as the most likely to produce results better than random guessing.

Text analysis, and more specifically sentiment detection could provide an insight into investor and general public opinion on a company and its stock price on a large scale. This insight could provide more information for use in analysis techniques similar to those currently supported by technical analysis. This could be a promising method for determining the relationship between human evaluations and stock price apart from the apparent underlying values of companies uncovered by fundamental analysis.

2. Related Work

There has been a significant amount of research into text analysis, including sentiment analysis, as well as some interest in utilizing these tools for prediction through Twitter, however up until now these projects have primarily worked with text analysis and sentiment prediction more generally. This is one of the unique difficulties of the problem of detecting investor sentiment on Twitter. Since tweets expressing clear sentiment about a stock can look either objective or simply noisy to general models. For example one collected tweet reads "$MSFT bullish http://t.co/u1Z74OGRDS" which has little natural meaning, however in the context of the jargon particular to securities markets, this tweet expresses clear positive sentiment towards Microsoft Corporation. Such difficulties necessitate the construction of a sentiment classifier particular to this field of study. General models such as the Stanford NLP sentiment classifier discussed in the introduction, however, can still be immensely valuable in providing a basic framework for a context specific classifier.

2.1. Textual Representation

There are three primary means of representing text in statistical textual analysis. These are n-gram, vector space modeling, and character streams. The first of these techniques, n-gram representations, has been around for decades and provides the simplest most straightforward method of representing text based on simple word or character sequence counts. Vector space models are a more complex and far more recent development with the most popular implementation, "word2vec", having been created within the last few years. Finally character stream techniques are the most recent development in the field with the first viable model having been published mere months before this writing. As such this final technique, though an extremely promising avenue for progress in the field, has been omitted from this research. Its likely impact however is significant enough to justify inclusion in any overview of textual representation techniques.

N-gram representations are based on simple character or word sequence counts. In these techniques a full corpus of related text is parsed, and every appearing character or word sequence of length n is extracted to form a dictionary of words and phrases. For example the text "the quick brown fox jumps over the lazy dog" has the following 5-gram word features: "the quick brown fox jumps", "quick brown fox jumps over", “brown fox jumps over the”, “fox jumps over the lazy”, and “jumps over the lazy dog”. Similarly the text “g2g ttyl” has the following 5-gram character features: “g2g t”, “2g tt”, “g tty”, and “ttyl”. Every text in the corpus can then be easily marked as a vector based on a simple count of the number of times each phrase in the dictionary occurs in the text. The main advantages of this technique are its simplicity, and flexibility to specifically match the corpus of text being studied [6].

Vector space techniques require a substantial set of text cleaned so as to include only the words of the language. The spatial relationship between the words is then analyzed as described by Mikolov et. al. [7]. Once every word of the language has been mapped to a unique vector, collections of words can be aggregated using entry wise summation and normalization yielding a vector for any given collection of words [8]. The cosine similarity between the vectors of words in such a representation demonstrates a significant carryover in word meaning [9]. For example the vector for “queen” can be calculated from the vectors for “king”, “man”, and “woman” as follows: “king” – “man” + “woman” = “queen”. This sustained relationship between word concepts makes vector space models very attractive for textual analysis. The primary negative consequence of vector space representations is their need to form a global representation across an entire language which may cause difficulty in interpreting the language particular to a specialized field where words might be used with slightly altered meanings and relationships.

Character stream modeling requires only a dictionary of valid characters to learn from a text. This enables it to be language agnostic to the point that the same algorithm can work effectively on languages as diverse as English and Chinese when provided adequate training data. This is a huge advancement over previous methods which often maintain rigid requirements for the language they are designed to model. So far this technique demonstrates superior predictive power to other methods albeit at the cost of increased computational complexity [10].

2.2. Sentiment Analysis on Twitter

In studying sentiment analysis on noisy and biased data, it was found that a multi-level classification model can provide more robust, and accurate predictions in difficult data sets [11]. In this model tweets are first classified based on objectivity versus subjectivity. This can remove a lot of the noisy advertising data that simply states facts and separately uncovers tweets that are more likely to express true sentiment. These can then be classified with greater accuracy since the vast majority of tweets lacking sentiment that would have caused difficulties have been eliminated and effectively classified as neutral. This could be particularly useful in evaluating tweets about stocks since both subjective and objective tweets can provide information about the market, but would require separate processing methods.

There have also been several deep learning approaches to sentiment classification on twitter that have been specialized to account for the relatively limited data available in a text with a maximum of one-hundred forty characters. These studies have found that a combination of specially chosen metadata and textual features, along with more traditional analyses such as n-grams can provide a more accurate classification model than simple features alone. [11][12] In particular targeting the appearance of words in capital letters, emoticons, and elongated words such as “whaaaaaat” or “coooooool” can significantly improve classification. Any approach to sentiment analysis utilizing deep neural networks deserves consideration considering the outstanding performance such methods have demonstrated in recent years.

2.3. Predicting the Behavior of a Population Based on Sentiment

There is some precedent for using aggregated sentiment analysis from Twitter data to make predictions about the behavior of a population. Previous research has demonstrated a significant connection between the overall sentiment of Twitter users towards new movie releases and the box office sales figures [13][14]. Bollen et. al. found that the overall mood of an unfiltered Twitter stream can predict movements in the Dow Jones Industrial Average as a whole [15]. Similarly the movement of markets of highly related companies has been predicted using an aggregate of the sentiment of companies in that market [16]. It is then reasonable to believe that the sentiment related to a particular company within the Dow Jones Industrial Average might be indicative of its future short term price movements.

3. Data and Methods

In order to perform a comparison between stock price changes and Twitter sentiment it was necessary to collect data on both trading values and tweets related to companies with timestamps to properly match the two into a consistent stream. In order to narrow down the range of data needing collection this study focused on the thirty companies of the Dow Jones Industrial Average. Data was collected from both sources over a period of several months from November 2014 through March 2015 utilizing specialized APIs. Only the data collected between February 6th, and February 18th was fully evaluated due to computational limitations and gaps in tweet data caused by throttling.

3.1. Twitter and Stock Price Data Collection

Twitter provides its own API through which developers may obtain limited streams of live tweets. This interface was used through the Twitter4j java library to filter all available live tweets for any containing any complete company name, or ticker symbol on the Dow Jones Industrial Average. All tweets matching the filter were saved along with all available metadata including timestamp, sender, geotagging, retweet status etc. The number of tweets for each company on each day is recorded in **table 1**. Clearly some companies are far more commonly mentioned on Twitter. Similarly the Yahoo! Finance stock API was used to collect stock price data on the stocks of the Dow Jones Industrial Average. Every five minutes the API was queried for bid, ask, and last trade prices. This information was recorded for each stock along with the timestamp identifying precisely when the information was collected.

Because all of this information was collected in a real time streaming environment with very brief time windows, and no modifications to the data, this approach lends itself well to the type of moment by moment analysis that must be conducted for technical stock analysis.

Table 1: Number of tweets collected for each company by day

	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9	Day 10	Day 11	Day 12
AXP	702	1014	745	691	566	493	657	620	675	642	832	608
BA	2611	2646	2147	2170	2055	1990	2071	2239	2251	2416	2337	1896
CSCO	2576	3060	2061	1888	2389	1885	1887	2403	2081	2392	2020	2139
CVX	1247	1112	1155	1270	1786	1325	1382	1251	1524	1106	1308	1205
DD	57	56	67	76	74	58	77	54	79	62	75	60
DIS	63479	56041	55236	62317	55266	59685	58017	62359	66750	57061	61879	60274
GE	270	195	362	247	346	273	218	256	295	256	269	205
GS	1014	993	1081	884	967	872	712	831	1129	991	1094	689
HD	1317	1367	1302	1351	1027	1370	1253	1500	1375	1139	1177	1056
IBM	79	68	71	76	56	76	73	70	59	71	85	75
INTC	7393	6893	6134	7288	6682	6937	6391	7542	7469	6852	7285	7210
JNJ	36	48	72	49	50	42	35	38	49	41	61	26
JPM	374	494	380	450	412	486	463	358	402	418	541	409
KO	4662	4384	4130	4329	4425	3883	4347	5143	4134	3615	5584	4034
MCD	7932	8826	7386	7200	6945	7518	7315	7523	7829	7837	8340	9049
MMM	1195	1432	930	1100	1115	1081	1005	1282	1513	1226	1797	1605
MRK	213	171	224	238	215	224	233	216	203	194	216	217
MSFT	26913	26819	23463	28622	26520	25017	24863	27484	28378	25847	28454	26734
NKE	43457	40396	38905	47152	43407	42238	40601	44277	43570	40947	44277	46695
PFE	441	336	337	408	436	434	338	372	376	401	366	392
PG	21	19	26	21	14	29	21	34	25	21	16	17
TRV	25	18	16	41	27	32	20	11	46	20	16	19
T	32	55	34	77	34	36	28	56	65	46	42	25
UNH	110	97	93	125	102	97	94	125	110	101	112	107
UTX	36	31	25	48	25	32	32	49	38	30	42	48
VZ	5078	4716	4748	5973	4802	4828	4739	5234	5277	4359	5286	4937
V	4014	3725	3346	4287	4397	3357	3476	3418	3946	3638	3700	3898
WMT	878	730	721	882	687	798	867	973	812	838	790	818
XOM	486	537	578	691	582	567	721	821	398	563	582	524

3.2. Sentiment Analysis of Collected Tweets

After collecting raw text data from it was also necessary to compute the sentiment value for each tweet since it is this sentiment which may relate closely to the stock price. Initially the Stanford NLP Sentiment Classifier was used to predict the sentiment of each tweet. In order to evaluate the accuracy of these predictions it was necessary to prepare a set of tweets labeled with true sentiment values. This was done manually to ensure accuracy, and as such the labeled set consisted of only one thousand tweets. Due to the nature of the parser and its primary training on movie reviews and newspaper articles, it was particularly inadequate for the task performing with approximately 30% accuracy.

In response to this I constructed my own classification models, one using n-gram, and the other “word2vec” textual representation techniques to preprocess raw text before using a standard random forest model for classification. Each of these models performed with accuracy between 60% and 70% on the labeled data set, an acceptably high level of accuracy for textual sentiment analysis on such short texts.

3.3. Correlation Analysis on Stock Price and Tweet Sentiment

Since there are only two variables involved for each company, namely sentiment and price, Pearson’s Correlation Coefficient can readily demonstrate a connection between the two. In order to calculate the correlation between these values, sentiment values over five minute increments had to be aggregated. This allows a pairing of the sentiment over a five minute period with the value of the stock after five minutes. Unfortunately these numbers are of very different kinds. The sentiment value takes into account only the previous five minutes, while the stock value at specific moment takes into account everything that has occurred before hand. In order to create a more proper comparison two techniques were used. The first changes price values to account only for the last five minutes by using the price change since the previous measurement, while the second uses a running total of sentiment to help sentiment values aggregate beyond their five minute periods. Each of these techniques still leaves values in very different ranges, to correct this both sequences were normalized to fall between zero and one. Once this normalization was complete it was possible to calculate correlations between these series for each company. Unfortunately there remained a significant amount of noise because of the limited data available for any given five minute interval, and the variability of readings. To correct this a moving average was used with a window length of one day, and a step size of one hour. This significantly smoothed both curves and made results far more readable. This made it possible to plot these variables over the recorded time range to visually evaluate the relationship between price and sentiment for each company. A sampling of these calculations and charts may be found it the next section.

4. Results and Conclusions

4.1. Sentiment Classification Results

This section describes the results obtained through the methodologies described in the preceding section. The foundational problem to this study was the sentiment classification which was utilized by all subsequent testing methods. The confusion matrices of the classification using n-gram and vector space representations are shown in **tables 2-3**. Values on the top left to bottom right diagonal indicated correct predictions, while all off diagonal values indicate what errors the classifier is making in prediction. The sum of the diagonal values divided by the sum of all values in the table gives an accuracy of 68.5% for n-gram, and 63.4% for word2vec. This indicates that in this particular data set, the value of a domain specific vocabulary training outweighs the benefit of more accurate representation of the relationship between words for sentiment classification. While not outstanding for sentiment prediction in general, these figures are passable for unfiltered data in such a domain specific context.

Table 2: n-gram sentiment prediction

		Predicted Label		
		positive	neutral	negative
True Label	positive	179	126	18
	neutral	70	461	14
	negative	17	70	45

Table 3: word2vec sentiment prediction

		Predicted Label		
		positive	neutral	negative
True Label	positive	137	173	13
	neutral	47	479	16
	negative	18	98	16

These matrices show that using the n-gram representation predictions on positive, neutral, and negative tweets had accuracies of 55.4%, 84.6%, and 34% respectively. Similarly the same accuracies using the word2vec representation were 42.4%, 88.4%, and 12.1%. Preference for neutral sentiment is due to the overall probability of a given tweet having neutral sentiment. Since most tweets are neutral the classification model errs towards neutral predictions. This leads to a prediction biased towards positive since positive prediction is higher than negative prediction, however given sufficient training data this would not be an issue as this bias is caused by the occurrence of more positive tweets than negative ones in the sample data. The predictions remain in proportion with the true labels of the training data.

4.2. Price and Sentiment Correlation Analysis Results

Results of the correlation analysis were very mixed across companies, as may be seen in **table 3**. Some companies such as Cisco Systems (CSCO) and Goldman Sachs (GS) showed a strong negative correlation of -0.90 between sentiment and price, while others such as Walmart (WMT) and Microsoft (MSFT) showed strong positive correlations of 0.85. These results are clearly visible in the charts for the selected companies shown in **figures 1-4**. From these it appears that in some cases there very well may be a connection between sentiment and price, but that connection may not always be the same.

Table 4: Pearson correlation coefficients for every company for each representation method

Company (symbol)	word2vec representation correlation	n-gram representation correlation		Company (symbol)	word2vec representation correlation	n-gram representation correlation
WMT	0.848094	0.847447		XOM	-0.41886	-0.45982
MSFT	0.856233	0.844525		MRK	-0.4751	-0.51843
UTX	0.841241	0.839601		KO	-0.53817	-0.52242
UNH	0.796927	0.815517		MCD	-0.52038	-0.52473
GE	0.787573	0.796093		INTC	-0.51058	-0.62937
IBM	0.902387	0.745618		T	-0.69758	-0.68944
DD	0.725684	0.698454		MMM	-0.73474	-0.71081
AXP	0.58826	0.668651		NKE	-0.73293	-0.73339
PFE	0.796572	0.598452		BA	-0.73247	-0.75375
CVX	0.40015	0.424398		JPM	-0.79279	-0.75673
TRV	0.489148	0.38869		HD	-0.78832	-0.78254
VZ	0.376637	0.380748		DIS	-0.86191	-0.81751
JNJ	0.334807	0.286778		CSCO	-0.91567	-0.89665
PG	-0.28735	-0.33896		GS	-0.90957	-0.90044
V	-0.39149	-0.37326				

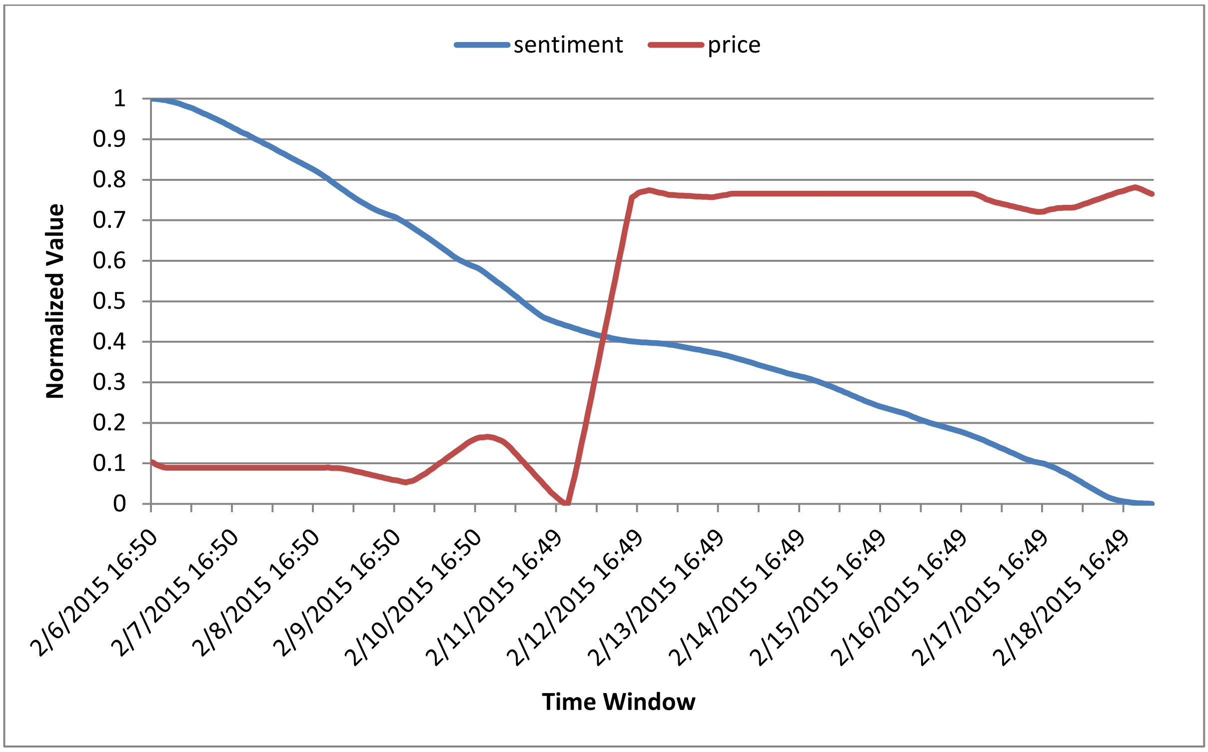

Figure 1: Cisco Systems price-sentiment Correlation (negative correlation)

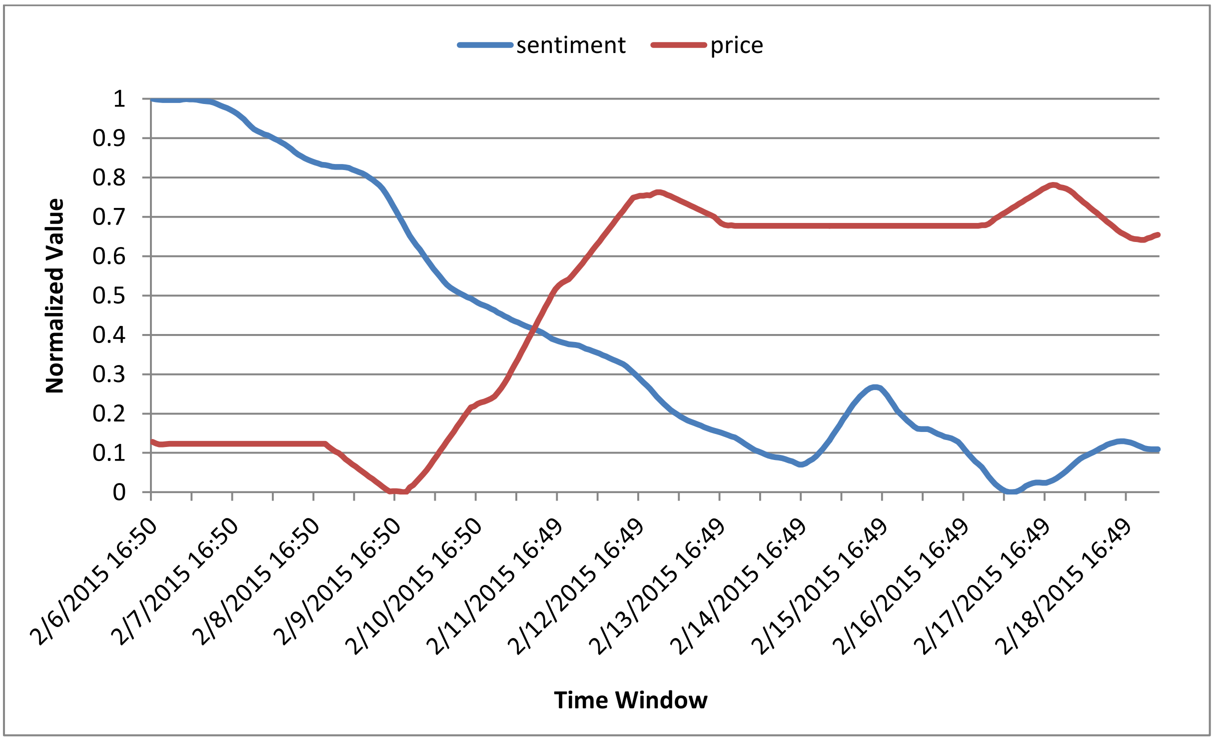

Figure 2: Goldman Sachs price-sentiment Correlation (negative correlation)

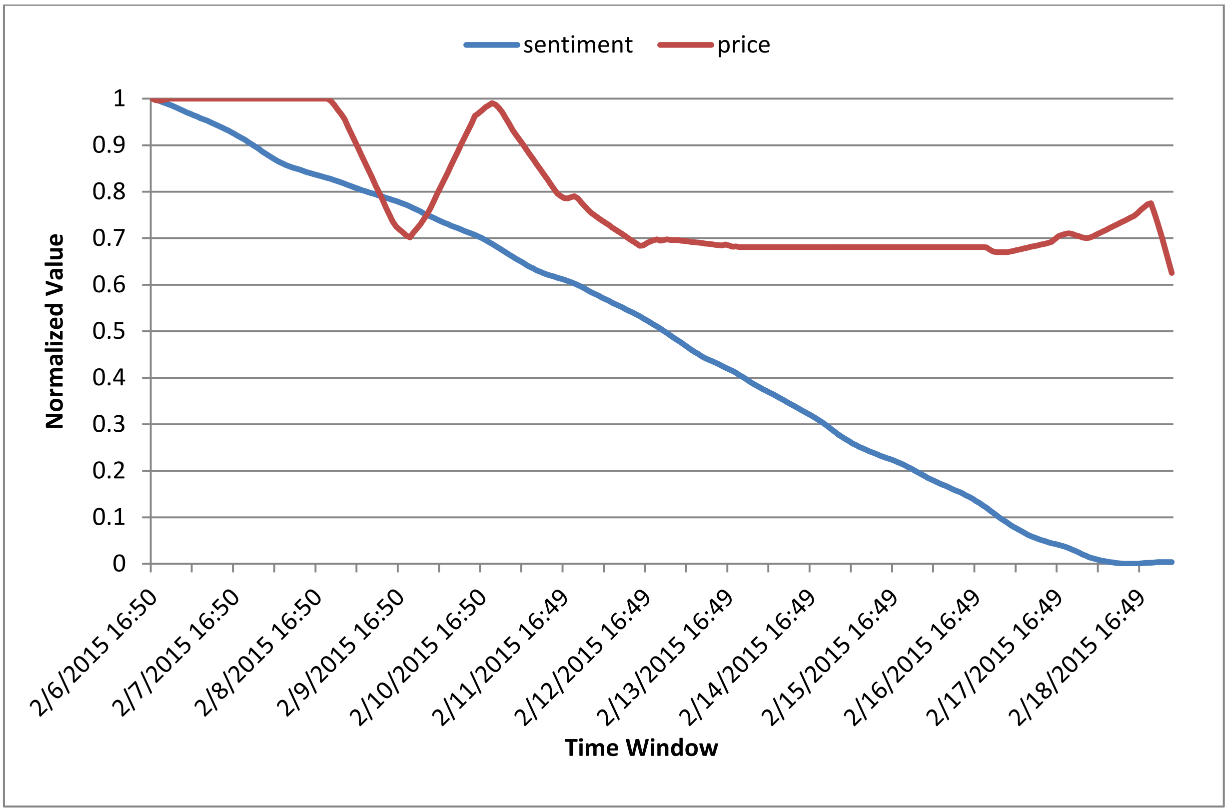

Figure 3: Walmart price-sentiment Correlation (positive correlation)

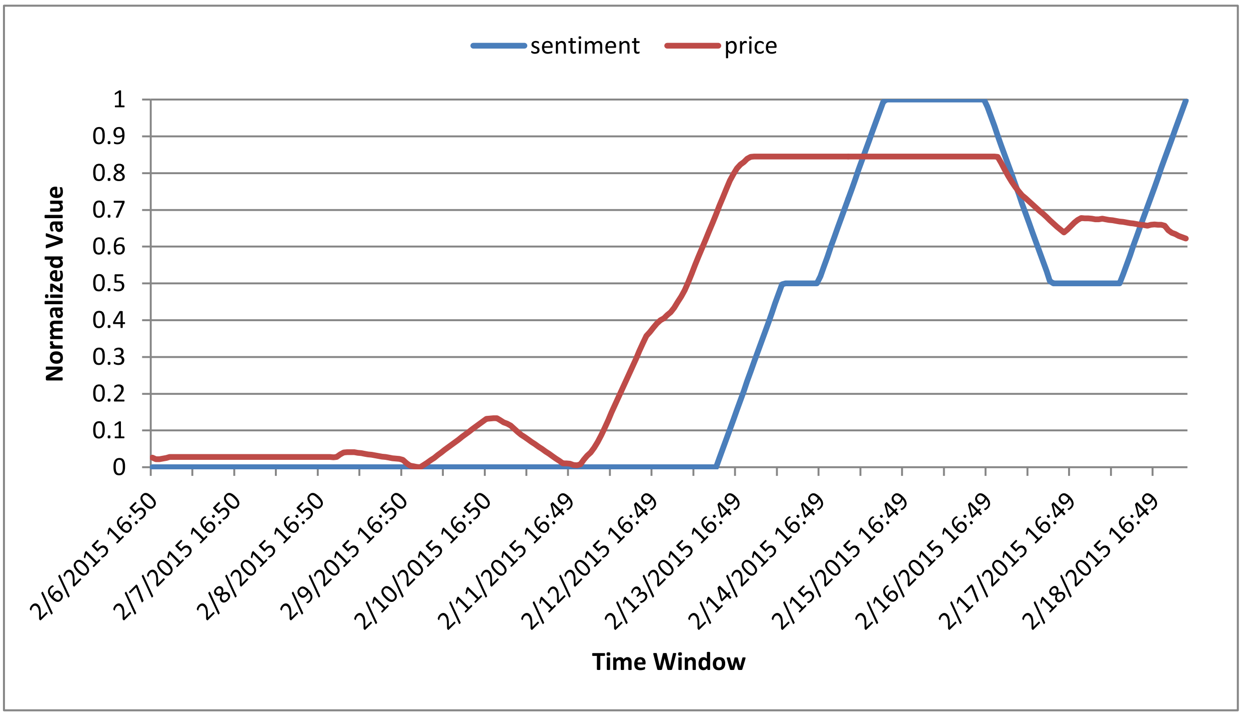

Figure 4: Microsoft price-sentiment Correlation (positive correlation)

4.3. Conclusions

The correlation between sentiment and stock price varies considerably across various companies with some having strong positive correlations, some having little correlation, and some having strong negative correlations. While they do not demonstrate a clear uniform connection between sentiment and price in single stock analysis across all companies, this does not preclude the possibility that different sources and expressions of sentiment might have different conflicting influences on price. This would support the findings of very different levels of correlation between sentiment and price across companies. Based on promising results in sentiment to price correlations on company groups from previous studies and on the strong correlation between sentiment and price for certain individual companies in this study, I believe that splitting different sources and expressions of sentiment including and more precise measures of sentiment target evaluation could uncover clear correlations between sentiment and stock price.

References

[1] "Top 20 Facebook Statistics" Zephoria Inc. February 10, 2015. Accessed March 10, 2015.

[2] Yang, Bishan, and Claire Cardie. "Context-aware Learning for Sentence-level Sentiment Analysis with Posterior Regularization." *Proceedings of ACL*, 2014.

[3] Socher, Richard, Alex Perelygin, Jean Wu, Jason Chuang, Christopher Manning, Andrew Ng, and Christopher Potts. "Recursive Deep Models for Semantic Compositionality Over a Sentiment Tree-bank." *Conference on Empirical Methods in Natural Language Processing*, 2013.

[4] Lev, Baruch, and S. Ramu Thiagarajan. "Fundamental Information Analysis." *Journal of Accounting Research*31, no. 2 (1993): 190-215.

[5] Wong, Wing-Keung, Meher Manzur, and Boon-Kiat Chew. "How Rewarding Is Technical Analysis? Evidence From Singapore Stock Market." *Applied Financial Economics* 13, no. 7 (2003): 543-51.

[6] Cavnar, William, and John Trenkle. "N-Gram-Based Text Categorization." *Proceedings of Annual Symposium on Document Analysis and Information Retrieval*, 1994.

[7] Mikolov, Tomas, Kai Chen, Greg Corrado, and Jeffrey Dean. "Efficient Estimation of Word Representations in Vector Space." *Computation and Language*, 2013.

[8] Mikolov, Tomas, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffery Dean. "Distributed Representations of Words and Phrases and Their Compositionality." *Computation and Language*, 2013.

[9] Mikolov, Tomas. "Linguistic Regularities in Continuous Space Word Representations." *Proceedings of North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2013.

[10] Zhang, Xiang, and Yann LeCun. "Text Understanding from Scratch." *Computation and Language*, 2015.

[11] Barbosa, Luciano, and Junlan Feng. "Robust Sentiment Detection on Twitter from Biased and Noisy Data."*Proceedings of the 23rd International Conference on Computational Linguistics*, 2010, 36-44.

[12] Tang, Duyu, Furu Wei, Bing Qin, Ting Liu, and Ming Zhou. "Coooolll: A Deep Learning System for Twitter Sentiment Classification." *Proceedings of the 8th International Workshop on Semantic Evaluation*, 2014, 208-12.

[13] Rui, Huaxia, Yizao Liu, and Andrew Whinston. "Whose and What Chatter Matters? The Effect of Tweets on Movie Sales." *Decision Support Systems* 56, no. 4 (2013): 863-70.

[14] Mesnil, Gregoire, Tomas Mikolov, Marc'Aurelio Ranzato, and Yoshua Bengio. "Ensemble of Generative and Discriminative Techniques for Sentiment Analysis of Movie Reviews." *Computation and Language*, 2014.

[15] Bollen, Johan, Huina Mao, and Xiaojun Zeng. "Twitter Mood Predicts The Stock Market." *Journal of Computational Science* 2, no. 1 (2011): 1-8.

[16] Si, Jianfeng, Arjun Mukherjee, Bing Liu, Sinno Pan, Qing Li, and Huayi Li. "Exploiting Social Relations and Sentiment for Stock Prediction." *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 2014, 1139-145.