# Predicting Big Movers Based on Online Stock Forum Sentiment Analysis

*Full Paper*

**Mohammad Al-Ramahi**
Dakota State University
Madison, SD
maabdel-rahman@pluto.dsu.edu

**Jun Liu**
Dakota State University
Madison, SD
jun.liu@dsu.edu

**Omar El-Gayar**
Dakota State University
Madison, SD
Omar.El-Gayar@dsu.edu

**Yenling Chang**
Dakota State University
Madison, SD
yenling.chang@dsu.edu

## Abstract

While social media sentiment has been proved to have predictive value for stock indices, it is intriguing to investigate if it is useful for predicting price changes for individual stocks. We focus on a special kind of stocks, big movers, i.e., stocks that undergo a drastic one-day price change, and a special kind of social media, online stock discussion forums. Based on an empirical study, our research shows that discussions during the days lead up to the big one-day price change do contain sentiments that can be used to predict big movers. The findings of our research have theoretical implications for future research on social media sentiment and practical implications for developing stock investment strategies.

### Keywords

Big movers, text mining, sentiment analysis

## Introduction

Behavioral theories posit that investors are not "rational" but "normal" and that systematic biases in their beliefs induce them to trade on non-fundamental information, called "sentiment" (Li et al. 2010). Researchers have used both direct and indirect sentiment measures and showed that investor sentiment exhibits predictive power for stock returns. Direct sentiment measures are derived from surveys. For instance, the American Association of Individual Investors (AAII) has conducted a sentiment survey among its members since July 1987. The survey results have been used in existing research (e.g., Brown and Cliff (2004), Fisher and Statman (2004)) to show the correlation between stock returns and investor sentiment. Indirect sentiment measures refer to financial variables that are related to sentiment. As an example, Baker and Wurgler (2006) have constructed an investor sentiment index based on indirect measures such as trading volume, the first-day returns on IPOs, etc.

More recently, with the advent of social media and Web 2.0, user generated content provides a new source of investor sentiment. Millions of investors are posting their opinions regarding individual stocks and the overall stock market on social media platforms such as Yahoo Finance, Raging Bulls, Seeking Alpha, and StockTwits.com, just to name a few. User generated content in social media enriches investors' ability in making better investing decisions by allowing investors to monitor the thought process and decision makings of others. It also gives researchers an unprecedented opportunity to utilize the large-scale data to investigate the impact of social media sentiment. Recent studies such as (Bollen and Mao 2011; Zhang et al. 2012) have found that social media sentiment is correlated with the movement of stock indices such as Dow Jones, NASDAQ and S&P 500. For example, Bollen and Mao (2011) identify a correlation between

Twitter mood and DJIA values, and they admit causality is unknown. Scholars, however, have provided little evidence that these sentiments play any significant role in predicting individual stock price movements (Antweiler and Frank 2004; Das and Chen 2007; Tumarkin and Whitelaw 2001). Our study attempts to address this research gap by focusing on a specific kind of stocks called "big movers". We define big movers as the stocks that have at least 15% one day price change. Oppositely, non-big movers refer to those stock that posted low percentage (<3%) one-day price changes. Unlike most stocks that normally post low percentage daily price changes, big movers signify outstanding short-term returns. They can be categorized into "big gainers" and "big losers". Predicting big movers based on sentiment therefore has significant practical implications for investors. We examine the predictive value of sentiment of one type of social media, namely online stock discussion forums for big movers. Of the stock discussion forums on the Internet, we focus on Yahoo Finance forum. Yahoo Finance hosts one of the largest and most popular online communities, with boards for over 6000 stocks.

To assess the predictive value of sentiments revealed in Yahoo finance discussion forum for big movers, we put forward the following research questions:

1. How well can stock discussion forum sentiment identify big movers from non-big movers?

2. Can stock discussion forum sentiment be used to identify big gainers vs big losers effectively?

We propose to extract relevant features in building robust models for predicting big movers based on sentiment analysis of stock forum discussions. According to (Feldman 2013), the sentiment lexicon is the most crucial resource for most sentiment analysis algorithms. We develop our own sentiment lexicon by extracting a set of terms that indicate bullish or bearish sentiments from Yahoo stock discussions. Using the terms defined in the lexicon as well as those defined in (Loughran and McDonald 2011) as features, we develop classification models that effectively distinguish big gainers and big losers from non-big movers. Overall, our study shows that online discussion forum sentiments do contain valuable information for investing decision making regarding individual stocks.

## Related work

Research that is related to ours includes those that examine the relationship between social media sentiment and stock prices (e.g., Antweiler and Frank 2004; Sprenger et al. 2014; Sul et al. 2014). In recent years, both researchers and practitioners have increasingly looked towards sentiment analysis as a tool for predicting stock prices. Yu et al. (2013) showed that overall social media sentiment has a stronger impact on firm stock performance than conventional media. Also, there are existing studies that have used online message boards as a mean to aggregate investor sentiment. Antweiler and Frank (2004) studied the predictive power of online message boards for the stock market by analyzing 1.5 million messages from Yahoo! Finance and Raging Bull. Applying sentiment analysis, they found that the number of messages is a predictor for stock turnover. However, their model's performance would not deliver a significant return on investment comparable with plausible transaction costs. Das et al. (2005) found that discussion forum sentiment follows stock price returns. Sentiment does not apparently predict returns, but returns drive sentiment.

In recently year, twitter has been recognized as a source of investor sentiment. Bollen and Mao (2011) aggregated "tweets" from Twitter as a whole, choosing not to focus on stock market specific "tweets," and examined the "mood" states of Twitter users and corresponding stock market movements. They showed the collective "mood" of Twitter users successfully predicted the upward and downward movement of the stock market 86.7% of the time. Sprenger et al. (2014) used a more direct approach to examine the relation between Twitter messages and stock market movement by filtering out all non-market related "tweets". Using the S&P 100 as the market index for analysis, their research suggests that public sentiment conveyed through StockTwits.com aligns itself with the movement of the S&P 100 and is positively related to the volume of trading. Sul et al. (2014) analyzed the cumulative emotional sentiments of Twitter posts about firms in the S&P 500. The results revealed a significant relation between the cumulative emotional valence of "tweets" about a firm and its stock returns.

The existing studies show that online discussion forums and twitter provides means for aggregating investor sentiments that are valuable for predicting stock indices changes. One issue that is baffling to researchers is that little research has succeeded in finding significant correlation between sentiment and individual stock price movements. Das and Chen (2007) attribute this to the large amount of noise in stock market message boards as well as the lack of market power that many investors participating in online message boards have. In our research, we focus on big movers. While social media sentiment may be ineffective for most stocks that post small or moderate price change, we attempt to investigate if the sentiment is useful for stocks that undergo drastic price changes.

# Data and Methodology

## *Data*

We collected historical prices of all stocks in NYSE and NASDAQ from March 2012 to March 2014 from Yahoo Finance[1]. Among the stocks, we identified big movers with one day price change greater than or equal to 15%. For each of the big movers, we collected five days' discussions on its Yahoo Finance discussion board before the date where the drastic price change occurred. We then removed stocks with less than 20 messages. Table 1 shows that we obtained 175 big movers including 98 big gainers and 77 big losers, and we collected on average 114.2 discussion messages for a big gainer and 103.7 messages for each big loser.

| Stock type | Count of stocks | Summary of discussions | | | | Stock price changes 5 days before the big move day |
|---|---|---|---|---|---|---|
| | | Count | Mean | Std Dev | Maximum | Mean |
| Big gainers | 98 | 11,190 | 114.2 | 152.2 | 777 | 4.12% |
| Big losers | 77 | 7,985 | 103.7 | 126.1 | 688 | 2.96% |
| Big movers | 175 | 19,175 | 109 | 139.2 | 733 | 3.54% |

**Table 1: Summary of big movers**

In order to compare the impact of forum sentiments on big movers vs non-big movers, we also collected data for some non-big movers. For each big gainer, i.e., a stock that moved drastically upward, we identified two stocks that had 0% - 3% gain on the day where the big move occurred. We also ensured that the non-big movers and the big mover belong to the same sector and industry. Similarly, for each big loser, we identified two stocks that belong to the same sector and industry and had less than 3% loss on the big moving day. We collected five days' discussions before the big moving day for each non-big mover and ensured that all non-big movers in our sample had at least 20 messages. Table 2 shows that we obtained 350 non-big movers. Unsurprisingly, during the five days leading up to the day the big move occurred, the big movers, including big gainers and big losers, have on average more discussion messages than their non-big moving counterparts.

| Stock type | Count of stocks | Summary of discussions | | | | Stock price changes 5 days before the non-big move day |
|---|---|---|---|---|---|---|
| | | Count | Mean | Std Dev | Maximum | Mean |
| Non-big gainers | 196 | 17,057 | 87.0 | 95.2 | 718 | 3.17% |
| Non-big losers | 154 | 13,651 | 88.7 | 89.9 | 486 | 2.79% |
| Non-big movers | 350 | 30,708 | 87.85 | 92.56 | 602 | 2.98% |

**Table 2: Summary of non-big movers**

Also, we captured the stock price behavior during the five days prior to the big move and non-big move day by computing the mean of the price change percentage during these days. Table 1 shows that the mean of the price behavior of big movers stocks is 3.54% compared with 2.98% of the corresponding non-big movers in Table 2.

[1] http://finance.yahoo.com/

### *Sentiments lexicon development*

The sentiment lexicon is the most important resource for sentiment analysis. Unfortunately, the effort in designing opinion lexicons adapted to financial terminology is scant. Loughran and McDonald (2011) manually created six word lists utilized in financial documents retrieved from the U.S. Securities and Exchange Commission web site over the years 1994 to 2008. However, according to our preliminary investigation, such lexicon is not targeted to online discussion text. We hence propose to develop our own sentiment lexicon that includes sentiment terms extracted from Yahoo stock discussion boards.

The poster on the Yahoo message boards has the ability to choose a sentiment disclosure of his/her recommendation for readers (i.e. "strong buy", "buy", "hold", "sell", and "strong sell") regarding a stock. Hence, some messages on Yahoo stock discussion boards have sentiments assigned by the posters. To develop our sentiment lexicon, we leveraged the user designated sentiments and randomly sampled 3,000 messages with "strong buy" or "buy" sentiments and 3,000 messages with "strong sell" or "sell" from discussion boards of 323 stocks. In order to avoid contamination of our training data (including the big movers and the non-big movers we selected for our study), the 323 stocks we used to construct our sentiment lexicon do not include those in our training data. We processed each message with sentiment by 1) transforming all upper case words in the message into lower case words, 2) splitting the message into a sequence of tokens using "non-letter" character as the splitting point, and 3) filtering the English stop-words from each message. We then counted the frequency of each word in all messages and kept those with frequency greater than 20. Next, we manually selected those words that we believe bring sentiment meanings in the finance context. As a result, we obtained 82 sentiments terms that were added to our sentiment lexicon. Examples of the sentiment terms in our lexicon include "overdone", "pop", "choppy", "rocketed", "explosive", "restore", etc.

### *Stock classification based on sentiment*

#### Models

We attempt to answer the research questions R1 and R2 with four predictive models: M1, M2, M3 and M4. In all of these four models, the target variable includes three different classes: 1) big gainer, 2) big loser, and 3) non-big mover. When collecting our samples, we obtained two non-big movers for each big gainer or big loser. Here we do not make a distinction between non-big movers that moved slightly upward and those that moved slightly downward. We created a common label "non-big movers" for all non-big movers in our dataset. Hence, our training data includes messages of 98 big gainers, 77 big losers and 350 non-big movers.

As described previously, we developed our own sentiment lexicon. In order to evaluate the effectiveness of using our lexicon in predicting big movers, we use the widely-used financial lexicon developed in (Loughran and McDonald 2011) as the baseline lexicon and compare our lexicon with it. We developed four models, each of which uses different features, to compare the effectiveness of our sentiment lexicon vs that of the lexicon developed by Loughran and McDonald.

In M1, we used the 82 words in our sentiment dictionary as features. M2 uses all words in the baseline lexicon as features. A problem with the baseline lexicon is that it contains a large number of sentiment words including 345 positive words and 2,329 negative words. If we use all of the words as features, such a large number of features can potentially cause the issue of overfitting. We hence propose M3. We first perform feature selection using the commonly used Chi-square (X2) method for feature selection. The Chi-square method evaluates features individually by measuring their Chi-square statistic with respect to the classes of the target variable. We selected only a number of features that have a Chi-square test score that is statistically significant at the 0.01 level. As a result, we obtained 49 sentiments terms from the baseline lexicon and used these terms as features of M3. These selected terms include "miss", "overpaid", "stabilized", "break", "collapsed", "disclaimer", "incredible", "suspended", "bailout", "sue", "prematurely", etc. Finally, to investigate if our lexicon used in conjunction with the lexicon achieve better performance, we combined the 82 sentiment words in our lexicon with the 49 words extracted from the baseline lexicon based on the Chi-square method, removed the 18 common words in both lexicons, and obtained 113 words. We then used these 113 words as features in M4.

After identifying the features for each of the four models, we constructed a feature vector including binary numbers for each stock in our training dataset. Each stock has a number of discussion messages, and we combined the messages. For each feature (i.e., a sentiment word) in a model, if the messages of a stock include the word, we assigned 1 to the stock's feature vector, and we assigned 0 otherwise. Since the target variable in our four models is categorical and includes more than two classes (including "big gainer", "big loser", and "non-big mover"), we propose to conduct multinomial logistic regression analysis.

### *Evaluation metrics*

To evaluate the predictive power of our models, we chose four evaluation metrics, precision, recall, accuracy, and AUC (area under the ROC curve). The precision metric evaluates the prediction accuracy by dividing the number of positive samples that correctly predicted as positive (TP) on the total number of both TP and those mistakenly classified as positive (FP). Note that the drawback of the precision is that it does not take into account those who are incorrectly classified as negative samples.

$$\text{Precision} = TP / (TP+FP) \tag{1}$$

On the other hand, the recall metric evaluates the prediction accuracy by dividing the number of TP on the total number of both TP and those are incorrectly classified as negative (FN).

$$\text{Recall} = TP / (TP+FN) \tag{2}$$

The accuracy metric measures the percentage of those correctly classified as positive or negative examples.

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN) \tag{3}$$

The last metric is the Area Under the ROC Curve (AUC). AUC is a standard measure for assessing the quality of predictive models (Bohannon et al. 2011). We conduct 10 fold cross-validation to evaluate the aforementioned four models.

## Results

Table 3 shows the performance results of the four models in predicting big gainers, big losers and non-big movers. M1 that uses 82 words in our own sentiment lexicon as the features achieved precision of 68%, recall of 72%, accuracy of 71%, and AUC of 78%. M2 with all words in the baseline lexicon as features performed poorly against M1 in all four metrics (precision of 56%, recall of 60%, accuracy of 60%, and AUC of 71%). Selecting 49 words from the baseline lexicon using the Chi-square method helped resolve overfitting and enhance the model fit. As a result, M3 achieved comparable (slightly worse only in recall) performance to M1. M4 demonstrated the best performance among all these four models in all four evaluation metrics (precision of 73%, recall of 76%, accuracy 75% and AUC of 80% respectively), which indicates that combining the sentiment terms in our own lexicon with those in the baseline lexicon constitutes the best set of the sentiment terms for predicting big movers.

| **Precision** | **Recall** | **Accuracy** | **ROC index** |
|---|---|---|---|
| Upward Big movers vs. Downward big movers vs. Non-big movers | | | |
| M1: Our lexicon | | | |
| 68% | 72% | 71% | 78% |
| M2: The whole baseline lexicon | | | |
| 56% | 60% | 60% | 71% |
| M3: The reduced baseline Lexicon with selected features | | | |
| 68% | 71% | 71% | 78% |
| M4: Our dictionary + The reduced baseline lexicon | | | |
| 73% | 76% | 75% | 80% |

**Table 3: The prediction results of the four models**

## Conclusion

In this study, we explored the extent to which sentiment revealed in Yahoo stock discussion forum can be used to distinguish big movers (i.e., stocks that have a drastic one-day price change) including big gainers and big losers from none-big movers. Our research makes the following contributions.

Theoretically, we contribute to the literature on impact of sentiment on stock price changes. As discussed previously, researchers have long been puzzled by the findings that social media sentiment has little predictive power for individual stocks. Hence, existing research on social media sentiment has been focusing on the correlation between the sentiment and stock indices. Our research sheds some light on the impact of social media sentiment on individual stocks. Not only do the big movers had the larger number of messages during the days leading up to the big move, but they also displayed stronger sentiments than non-big movers. Big stock price changes often occur due to official earnings announcements, news of lawsuits, rating upgrades or downgrades, FDA drug approvals or disapprovals, etc. Although social media sentiment may be of little predictive value for individual stocks on a normal day, our research shows that in anticipation of the "big event", discussions on online stock forums intensify, contain more sentiments, and these sentiments are indicative of the direction of pending stock price changes.

Methodologically, our research contributes to the area of sentiment analysis by developing a sentiment lexicon targeted to financial discussion forums. As discussed previously, sentiment lexicons are critical for sentiment analysis. Although there are sentiment lexicons developed for the finance domain, we lack those that are tailored to capture the sentiments of online stock discussions. In our research we developed a sentiment lexicon based on analyzing discussions on Yahoo Finance and proved its effective by comparing it with the well-known financial sentiment lexicon developed by (Loughran and McDonald 2011). Our research further shows that integrating terms contained in both lexicons helps improve the prediction performance.

Practically, our research has significant implications for stock investors. Big movers lead to big money. Identify big movers based on social media sentiment provides investors with a unique, appealing edge. Combining the insights obtained from social media sentiment with other technical and fundamental analysis methods, investors can develop trading strategies that grasp the trading opportunities provided by the big movers.

Future research can further refine the proposed lexicon by exploring ways to capture the intensity of the sentiments and possibly improve the classification performance for big movers. Moreover, while this research focused on Yahoo stock discussions, the generalizability of the approach can be further demonstrated and validated by exploring other social media venues. Another possibility for future research is leveraging the results of this research as an initial step for complementing technical and fundamental analysis with results from social media analysis to obtain a more holistic understanding of factors impacting changes in stock process (theoretical perspective) as well as improve the predictive potential of existing techniques (practical perspective).

## REFERENCES

Antweiler, W., and Frank, M. Z. 2004. "Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards," *The Journal of Finance* (59:3), pp. 1259-1294.

Baker, M., and Wurgler, J. 2006. "Investor Sentiment and the Cross-Section of Stock Returns," *The Journal of Finance* (61:4), pp. 1645-1680.

Bohannon, T., Hardin, M., Kelly, D., Lucas, B., and Walsh, S. 2011. "Applied Analytics Using Sas Enterprise Miner Course Notes." Cary, North Carolina, USA: SAS Institute Inc.

Bollen, J., and Mao, H. 2011. "Twitter Mood as a Stock Market Predictor," *Computer* (44:10), pp. 91-94.

Brown, G. W., and Cliff, M. T. 2004. "Investor Sentiment and the near-Term Stock Market," *Journal of Empirical Finance* (11:1), pp. 1-27.

Das, S., Martínez-Jerez, A., and Tufano, P. 2005. "Einformation: A Clinical Study of Investor Discussion and Sentiment," *Financial Management* (34:3), pp. 103-137.

Das, S. R., and Chen, M. Y. 2007. "Yahoo! For Amazon: Sentiment Extraction from Small Talk on the Web," *Management Science* (53:9), pp. 1375-1388.

Feldman, R. 2013. "Techniques and Applications for Sentiment Analysis," *Communications of the ACM* (56:4), pp. 82-89.

Fisher, K. L., and Statman, M. 2004. "Sentiment, Value, and Market-Timing," *The Journal of Investing* (13:3), pp. 10-21.

Li, X., Deng, X., Wang, F., and Dong, K. 2010. "Empirical Analysis: News Impact on Stock Prices Based on News Density," *Data Mining Workshops (ICDMW), 2010 IEEE International Conference on*: IEEE, pp. 585-592.

Loughran, T., and McDonald, B. 2011. "When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks," *The Journal of Finance* (66:1), pp. 35-65.

Sprenger, T. O., Tumasjan, A., Sandner, P. G., and Welpe, I. M. 2014. "Tweets and Trades: The Information Content of Stock Microblogs," *European Financial Management* (20:5), pp. 926-957.

Sul, H., Dennis, A. R., and Yuan, L. I. 2014. "Trading on Twitter: The Financial Information Content of Emotion in Social Media," *in Proceedings of the 47th Hawaii International Conference on System Sciences (HICSS)*, Waikoloa, HI: IEEE, pp. 806-815.

Tumarkin, R., and Whitelaw, R. F. 2001. "News or Noise? Internet Postings and Stock Prices," *Financial Analysts Journal* (57:3), pp. 41-51.

Yu, Y., Duan, W., and Cao, Q. 2013. "The Impact of Social and Conventional Media on Firm Equity Value: A Sentiment Analysis Approach," *Decision Support Systems* (55:4), pp. 919-926.

Zhang, X., Fuehres, H., and Gloor, P. A. 2012. "Predicting Asset Value through Twitter Buzz," in *Advances in Collective Intelligence 2011*. Springer, pp. 23-34.