

TweetTrader.net: Leveraging Crowd Wisdom in a Stock Microblogging Forum

Timm O. Sprenger

TUM School of Management
Chair for Strategy and Organization
Leopoldstrasse 139, 80804 Munich, Germany
timm.sprenger@gmail.com

Abstract

TweetTrader.net is a stock microblogging forum that leverages the wisdom of crowds to aggregate the information contained in stock-related tweets. Based on insights from academic research on stock microblogs, the application integrates inputs from text classification, user voting and a proprietary *Stock Game* in order to extract the sentiment (i.e., the bullishness) of online investors with respect to all publicly traded companies of the S&P 500. A demo of TweetTrader.net is available at <http://TweetTrader.net>.

Background of stock microblogging

Popularity of stock microblogs

Twitter has become a vibrant platform to exchange trading ideas and other stock-related information. Traders have adopted the convention of tagging stock-related messages (i.e. stock microblogs) with a dollar sign followed by the relevant company's ticker symbol (e.g., "\$AAPL" for tweets related to Apple Inc.). The business press describes the conversations on Twitter as "the modern version of traders shouting in the pits" (BusinessWeek 2009). There are investors who attribute their trading success to the information they find on social media websites and, moreover, financial professionals have developed Twitter-based trading systems to identify sentiment-based investment opportunities. As a result, the investor community has come to call Twitter and related third-party applications "a Bloomberg for the average guy" (BusinessWeek 2009).

Related academic research

A few recent studies suggest that the information content of general Twitter messages, including those without a specific reference to the stock market, may help predict

macroeconomic indicators such as the Index of Consumer Sentiment (O'Connor et al. 2010) or stock market indices such as the Dow Jones Industrial Average (Bollen, Mao, and Zeng 2010) and the S&P 500 (Zhang, Fuehres, and Gloor 2010). Regarding the information content of microblogs with respect to individual stocks, Sprenger and Welp (2010) find robust relationships between the sentiment of stock microblogs (i.e. their bullishness) and abnormal returns as well as message volume and trading volume. In addition, the authors offer an explanation for the mechanism leading to the efficient aggregation of information in microblogging forums by showing that users who provide above average investment advice are retweeted (i.e., quoted) more often and have more followers. In other words, retweets and followership may represent the Twittersphere's "currency" and provide it with a mechanism to weigh information.

The stock microblogging forum TweetTrader.net leverages the insights from this research and enables online investors to cut through the noise of thousands of daily messages. In contrast to a few related third-party applications such as StockTwits.com, which are limited to filtering the message stream by ticker symbol, TweetTrader.net taps the wisdom of crowds and aggregates the information in a meaningful fashion.

Features of TweetTrader.net

Tapping the wisdom of crowds

TweetTrader.net uses the Twitter Search API to provide a *Livestream* of all tweets related to S&P 500 stocks. Users have a choice between all tweets or a subset of selected indices or industries. Bullish words (e.g., buy, upgrade, or growth) are marked in green, bearish words (e.g., sold, downgrade, or decline) are marked in red. Embedded within these basic functionalities are three main features that tap the collective wisdom of thousands of online investors: automatic text classification, user-driven sentiment voting and a *Stock Game*.

Text classification. First and foremost, every stock-related tweet is automatically classified as either a buy (35.2% of all messages), hold (49.6%), or sell signal (15.2%) based on a manually coded training set of 2,000 tweets and the multinomial Naïve Bayesian classifier of the Weka machine learning package (Hall et al. 2009). 10-fold cross validation of the model classifies 64.2% of all messages correctly (for details regarding the classification approach, see Sprenger and Welpel 2010). The accuracy by class further validates the automatic classification of sentiment because the worst type of misclassification (i.e., of buy signals as sell signals and vice versa) occurs only rarely (less than 2% of all messages). Backtesting results illustrate that a trading strategy based on these signals would have earned profits of up to 15% over the first half of 2010.

Sentiment voting. The classification results demonstrate that the input from human coders can help correctly classify the sentiment of stock-related microblogs. Thus, next to every tweet users find two sentiment buttons (i.e., up and down) to vote the message as either a buy or a sell signal. This constant flow of manually coded messages can be used to update the automated classifier in real-time. Once users vote on a tweet, they can see how other online investors have rated the same piece of information. In addition, they have the opportunity to retweet this message and thereby give the information greater weight on the public timeline (e.g., because it will represent another data point for the text classifier).

Stock Game. The transformation into a retweet slightly alters the syntax of the original message. If a user votes a message as a buy (sell) signal, related ticker symbols are appended by a plus (minus) sign (e.g., “\$AAPL+” for a buy signal). These appended tags represent predictions for the end of day stock price relative to the stock price at the time of the prediction. A point is awarded for every correct prediction. Obviously, besides retweeting messages, users can simply enter *Stock Game* predictions directly by following the syntax either on the TweetTrader.net website or through any other Twitter client. TweetTrader.net keeps track of these predictions, evaluates them in real-time and shows a ranking of all participating users. This allows players to monitor their predictions and provides transparency for other stock microbloggers to identify and follow the best investment advisers in the Twittersphere.



Figure 1: *Livestream* with retweet function and sentiment voting

Aggregating stock-related information

The *Scoreboard* presents information related to a given stock. In particular, the buy and sell signals from the three above-mentioned inputs are combined into one sentiment signal, which shows how bullish or bearish (i.e. positive or negative) investors are with respect to a stock. In addition, because Sprenger and Welpel (2011) have shown that not only the sentiment but also the type of stock-related news events (e.g., financial issues vs. operations) affects subsequent returns, text classification methods, like those used to extract the sentiment, are employed to show the distribution of recently discussed topics for a stock (e.g., corporate governance, financial issues, or operations).

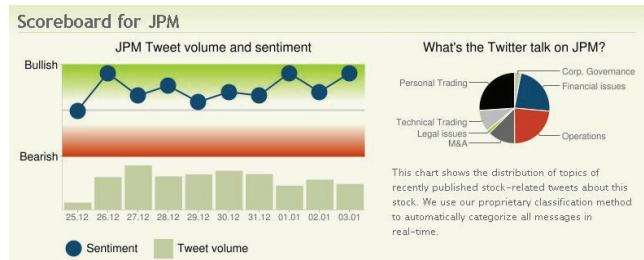


Figure 2: Sentiment and recently discussed topics (example JPM)

While TweetTrader.net is still in the early stages of development, this application shows promising ways to aggregate the collective wisdom of online investors in a meaningful fashion and enable stock microblogging forums to become to qualitative (e.g., textual) information what market mechanisms are to quantitative price signals.

References

- Bollen J., Mao H., and Zeng X.-J. 2010. Twitter mood predicts the stock market, *Arxiv working paper*, Indiana University.
- BusinessWeek 2009. StockTwits may change how you trade, *BusinessWeek (online edition)*, February 11.
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten Ian H., 2009, The WEKA data mining software: An update, *SIGKDD Explorations* 11, 10-18.
- O'Connor, B., Balasubramanian, R., Routledge, B. R., and & Smith, N. 2010. From Tweets to polls: Linking text sentiment to public opinion time series, In *AAAI International Conference on Weblogs and Social Media*, 122-129.
- Sprenger T., and Welpel, I. 2010. Tweets and Trades – the information content of stock microblogs, *SSRN working paper*, TUM School of Management.
- Sprenger T., and Welpel, I. 2011. News or Noise? The stock market reaction to different types of company-specific news events, *SSRN working paper*, TUM School of Management.
- Zhang, X., Fuehres, H., and Gloor, P. 2010. Predicting stock market indicators through Twitter – “I hope it is not as bad as I fear”, In *COIN Collaborative Innovations Networks Conference*, 1-8.