Sentiment analysis on tweets in a financial domain

Jasmina Smailović[1,2], Miha Grčar[1], Martin Žnidaršič[1]

[1] Dept of Knowledge Technologies, Jožef Stefan Institute, Ljubljana, Slovenia

[2] Jožef Stefan International Postgraduate School, Ljubljana, Slovenia

jasmina.smailovic@ijs.si

Abstract. This paper investigates whether sentiment analysis of public mood derived from large-scale Twitter feeds can be used to identify important events and predict movements of stock prices. We used the volume and sentiment polarity of Apple financial tweets to identify important events and predict future movements of Apple stock prices. Statistical analysis using the Granger causality test showed that we were able to predict the rise or fall in closing price of Apple stocks two days before the change happens.

Keywords: sentiment analysis, classification, Twitter, stock price prediction.

1 Introduction

Sentiment analysis or opinion mining [1] is a research area aimed at detecting the authors' attitude, polarity (positive or negative) or opinion about a given topic expressed in a document collection. In this paper we investigate whether sentiment analysis of public mood, derived from large-scale collections of daily posts from online microblogging service Twitter can predict movements of stock prices. Specifically, we analyse Apple financial tweets to identify important events and predict the movement of Apple stock prices.

Trying to determine the future revenue or stock value has been attracting great attention of numerous researches. Early research on this topic claimed that stock price movements do not follow any patterns or trends and past price movements cannot be used to predict the future ones [2]. Later studies, however, show the opposite [3][4]. Recent research indicates that the analysis of online texts such as blogs, web pages and social networks can predict trends of various economic phenomena. It was shown [5] that blog posts can be used to predict spikes in actual consumer purchase decisions. Sentiment analysis of weblog data was used to predict movie success [6]. Twitter posts were also shown to be useful when predicting box-office revenues of movies in advance of their release [7]. Furthermore, it has been shown [8] that the stock market itself is a direct measure of social mood. So, it is reasonable to expect that the analysis of public mood can be used to predict movement of stock market values. Moreover, Bollen et al. [9] show that changes in a specific public mood state can predict daily changes in the closing values of the Dow Jones Industrial Average index.

The paper is structured as follows: selection of data preprocessing settings for the SVM classifier is explained in Section 2, followed by an example in Section 3. Conclusions are given in Section 4.

2 Selection of data preprocessing settings for the SVM classifier

Here we describe how the most appropriate classifier for sentiment analysis of financial tweets was chosen. Three common approaches to sentiment analysis are: machine learning, lexicon-based methods and linguistic analysis. In this work we use the machine learning approach. In this approach, classification refers to a procedure for assigning a given piece of input data (instance) into one of a given number of categories (classes). In our case, input data is a tweet and it can be classified into one of two categories: positive or negative, which represent attitude of the tweet`s author. An instance is described by a vector of features (in our case, words and word pairs), also called attributes, which constitute a description of all known characteristics of the instance. An algorithm that implements classification is known as a classifier. Classification usually refers to a supervised procedure, i.e., a procedure that learns to classify new instances based on a model learnt from a training set of instances that have been properly labelled. For our training set we used a collection of 1,600,000 (800,000 positive and 800,000 negative) tweets collected by the Stanford University [10], where positive and negative emoticons were used as labels. For testing we used a set of manually labelled 177 negative and 182 positive tweets from the same source [10]. The SVM^{perf} classifier [11] was used for training and testing. It is an implementation of the *Support Vector Machine* machine learning algorithm. As attribute weights, we used TFIDF (term frequency–

inverse document frequency) which reflects how important a word is to a document in a collection or corpus. We explored the usage of unigrams, bigrams, replacement of usernames with a token, replacement of web links with a token, word appearance thresholds and removal of letter repetitions (e.g. 'looooove' is changed to 'love'). Table 1 summarizes the experimental results.

Table 1: Classifier performance evaluation for various preprocessing settings.

Maximum N gram length	Minimum word frequency	Replace usernames with a token	Replace web links with a token	Remove letter repetition	Accuracy	Precision/Recall
2	2	No	Yes	Yes	81.06%	81.32%/81.32%
2	2	No	No	Yes	78.83%	77.60%/81.87%
2	2	Yes	No	Yes	78.55%	75.86%/84.62%
2	2	Yes	Yes	Yes	78.27%	76.53%/82.42%
2	3	No	No	Yes	76.88%	77.97%/75.82%
1	2	No	No	Yes	76.32%	72.99%/84.62%

As it can be seen from the table, the best classifier is obtained by using both unigrams and bigrams, using words which appear at least two times in the corpus, with replacing links with a token and with removal of repeated letters.

3 Classifying financial tweets

Our main data resource for collecting financial Twitter posts is the Twitter API, i.e. Twitter Streaming and Search API. The Streaming API allows near-realtime access to various subsets of Twitter data while Search API returns tweets that match a specified query. By the informal Twitter conventions, the dollar-sign notation is used for discussing stock symbols. For example, $AAPL tag indicates that the user discusses Apple stocks. This convention simplifies the retrieval of financial tweets. We noticed that there are many tweets with similar content which are mainly a result of re-tweeting and spam. Twitter's re-tweet feature allows users to quickly post other users` messages. Spammers, on the other hand write nearly identical messages from different accounts. We employed the algorithm based on Jaccard similarity [12] to discard tweets that were detected as near duplicates. We analysed English posts that discussed Apple stocks in the period from March 11 to December 9, 2011. After pre-processing, 33,733 tweets were left and these were classified with the classifier described in Section 2. After classification, we count the number of positive and negative tweets for each day (Figure 1). Peaks show the days where people intensively talked about Apple. The analysis shows that these days correspond to important events.

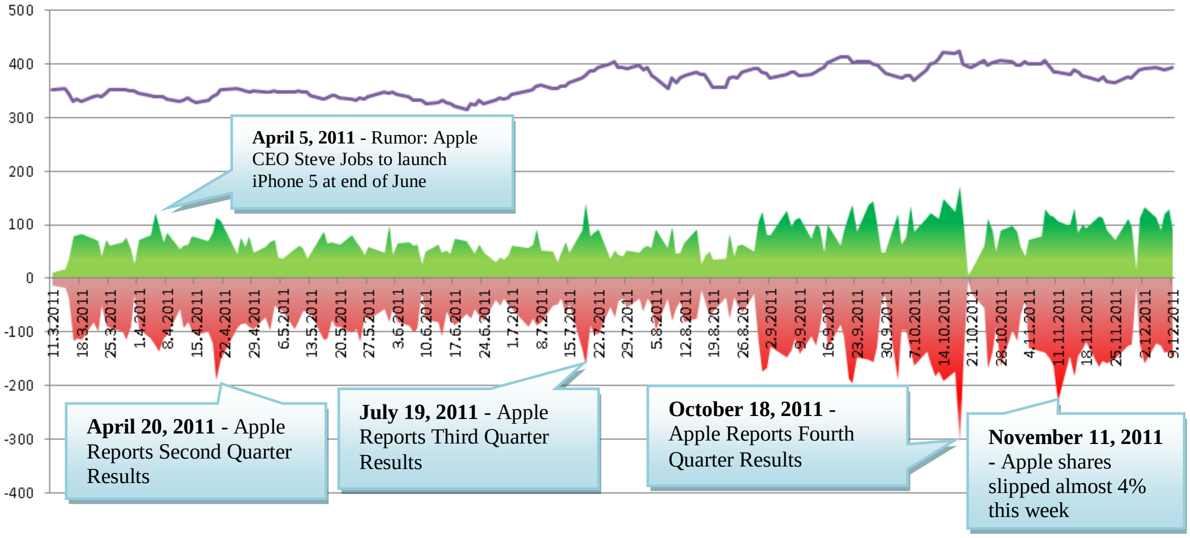

Figure 1: Number of positive (green), negative (red) tweet posts and closing price (violet) per day.

Next, we calculated the positive sentiment probability for each day. To enable the comparison of closing price and positive sentiment probability time series, we normalize them to z-scores. The z-score of time series *Xt*, is defined as:

$$Z_{Xt} = \frac{X_t - \bar{x}(X_{t\pm1})}{\sigma(X_{t\pm1})} \quad (1)$$

where $\bar{x}(X_{t\pm1})$ and $\sigma(X_{t\pm1})$ represent the mean and standard deviation of the time series within the period [t-1; t+1]. Next, we applied a statistical hypothesis test for determining whether positive sentiment probability time series is useful in forecasting the closing price. More specifically, we performed the Granger causality analysis [13] for the period between September 1 and December 8, 2011 as we notice that this is the period of big changes in the stock price when people also posted a large amount of messages. The Granger causality test (results shown in Table 2) indicates that positive sentiment probability could predict stock price movements, as we got a significant result (p-value < 0.1) in our dataset for a two day lag. This means that changes in values of positive sentiment probability could predict a similar rise or fall in closing price two days in advance.

Table 2: Statistical significance (p-values) of Granger causality correlation between positive sentiment probability and closing stock price.

Lag (days)	p-value
1	0.4855
2	**0.0565**
3	0.0872

4 Conclusions

Predicting future values of stock prices has always been an interesting task, commonly connected to the analysis of public mood. Various studies indicate that these kinds of analyses can be automated and can produce useful results as more and more personal opinions are made available online. In this paper, we investigated whether sentiment analysis of public mood derived from large-scale Twitter feeds can be used to identify important events and predict movements of stock prices. More specifically, Apple financial tweets were analysed, where our experiments showed that changes in values of positive sentiment probability with a delay of two days can predict a similar movement in the stock closing price. In the future, we plan to experiment with different datasets for training classifiers, analyse other companies‘ stocks and employ part of speech tagging in order to improve the classifier performance.

Acknowledgements

The work presented in this paper has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) within the context of the Project FIRST, Large scale information extraction and integration infrastructure for supporting financial decision making, under grant agreement n. 257928 and by the Slovenian Research Agency through the research program Knowledge Technologies under grant P2-0103.

References:

[1] P. Turney. Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. In *Proceedings of the Association for Computational Linguistics*, p. 417–424, 2002.

[2] E. Fama. Random Walks in Stock Market Prices. *Financial Analysts Journal*, 21(5): 55–59, 1965.

[3] K. C. Butler, S.J. Malaikah. Efficiency and inefficiency in thinly traded stock markets: Kuwait and Saudi Arabia. *Journal of Banking and Finance* 16, 197–210, 1992.

[4] M. Kavussanos, E. Dockery. A Multivariate test for stock market efficiency: The case of ASE. *Applied Financial Economics*, 11(5): 573-579, 2001.

[5] D. Gruhl, R. Guha, R. Kumar, J. Novak and A. Tomkins. The predictive power of online chatter. In *Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*, p. 78-87, New York, USA, 2005.

[6] G. Mishne and N. Glance. Predicting Movie Sales from Blogger Sentiment. In *AAAI Symposium on Computational Approaches to Analysing Weblogs AAAI-CAAW*, p. 155-158, 2006.

[7] S. Asur and B. A. Huberman. Predicting the Future with Social Media. In *Proceedings of the ACM International Conference on Web Intelligence*, arXiv:1003.5699v1, 2010.

[8] J. R. Nofsinger. Social Mood and Financial Economics. *Journal of Behavioral Finance*, 6(3): 144-160, 2005.

[9] J. Bollen, H. Mao, X. Zeng. Twitter mood predicts the stock market. *Journal of Computational Science*, 2(1): 1-8, 2011.

[10] A. Go, R. Bhayani, L. Huang. Twitter Sentiment Classification using Distant Supervision. *Association for Computational Linguistics*, p.30-38, 2009.

[11] T. Joachims. Training Linear SVMs in Linear Time. In *Proceedings of the ACM Conference on Knowledge Discovery and Data Mining (KDD),* 2006.

[12] A. Broder, S. Glassman, M. Manasse and G. Zweig. Syntactic Clustering of the Web. In *Proceedings of WWW6 and Computer Networks* 29:8-13, 1997.

[13] P. Wessa. Bivariate Granger Causality (v1.0.0) in Free Statistics Software (v1.1.23-r7), Office for Research Development and Education, http://www.wessa.net/rwasp_grangercausality.wasp/

For wider interest

From psychological research it is known that emotions are essential to rational thinking. Also, it has been shown that the stock market is a direct reflection of the social mood. On the other hand, more and more people make their opinions available publicly online, making it available for analysis. Can we expect that the analysis of public mood can identify important events and predict the movement of stock market values? Our preliminary studies indicate that the answer is – yes. We analysed the Apple financial Twitter posts that were collected in a 10 months period. We identified days when people intensively talked about Apple and consequently identified important events for this company. Next, we performed statistical analysis for the period of specific 3 months, which is the period of the main changes in the stock price, to determine whether we can predict future movement of Apple`s closing price. The test showed that we are able to predict the rise or fall in closing price two days before it occurs. This kind of analysis can also be applied to other domains. For example, it can be used for the assessment of products, prediction of purchase decisions, earnings and other similar phenomena.