Predictive Sentiment Analysis of Tweets: A Stock Market Application

Jasmina Smailović[1,2], Miha Grčar[1], Nada Lavrač[1], and Martin Žnidaršič[1]

[1] Jožef Stefan Institute, Jamova Cesta 39, 1000 Ljubljana, Slovenia
{jasmina.smailovic,miha.grcar,nada.lavrac,
martin.znidarsic}@ijs.si

[2] Jožef Stefan International Postgraduate School, Jamova cesta 39,
1000 Ljubljana, Slovenia

Abstract. The application addressed in this paper studies whether Twitter feeds, expressing public opinion concerning companies and their products, are a suitable data source for forecasting the movements in stock closing prices. We use the term predictive sentiment analysis to denote the approach in which sentiment analysis is used to predict the changes in the phenomenon of interest. In this paper, positive sentiment probability is proposed as a new indicator to be used in predictive sentiment analysis in finance. By using the Granger causality test we show that sentiment polarity (positive and negative sentiment) can indicate stock price movements a few days in advance. Finally, we adapted the Support Vector Machine classification mechanism to categorize tweets into three sentiment categories (positive, negative and neutral), resulting in improved predictive power of the classifier in the stock market application.

Keywords: stock market, Twitter, predictive sentiment analysis, sentiment classification, positive sentiment probability, Granger causality.

1 Introduction

Trying to determine future revenues or stock prices has attracted a lot of attention in numerous research areas. 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 future ones [1]. Later studies, however, show the opposite [2]. It has also been shown that emotions have an effect on rational thinking and social behavior [3] and that the stock market itself can be considered as a measure of social mood [4].

As more and more personal opinions are made available online, recent research indicates that analysis of online texts such as blogs, web pages and social networks can be useful for predicting different economic trends. The frequency of blog posts can be used to predict spikes in the actual consumer purchase quantity at online retailers [5]. Moreover, it was shown by Tong [6] that references to movies in newsgroups were correlated with their sales. Sentiment analysis of weblog data was used to predict movies' financial success [7]. Twitter[1] posts were also shown to be useful for predicting

[1] www.twitter.com

A. Holzinger and G. Pasi (Eds.): HCI-KDD 2013, LNCS 7947, pp. 77–88, 2013.

box-office revenues of movies before their release [8]. Thelwall et al. [9] analyzed events in Twitter and showed that popular events are associated with increases in average negative sentiment strength. Ruiz et al. [10] used time-constrained graphs to study the problem of correlating the Twitter micro-blogging activity with changes in stock prices and trading volumes. Bordino et al. [11] have shown that trading volumes of stocks traded in NASDAQ-100 are correlated with their query volumes (i.e., the number of users' requests submitted to search engines on the Internet). Gilbert and Karahalios [12] have found out that increases in expressions of anxiety, worry and fear in weblogs predict downward pressure on the S&P 500 index. Moreover, it was shown by Bollen et al. [13] that changes in a specific public mood dimension (i.e., calmness) can predict daily up and down changes in the closing values of the Dow Jones Industrial Average Index. In our preliminary work [14] we used the volume and sentiment polarity of Apple financial tweets to identify important events, as a step towards the prediction of future movements of Apple stock prices.

The paper follows a specific approach to analyzing stock price movements, contributing to the research area of sentiment analysis [15,6,16,17], which is aimed at detecting the authors' opinion about a given topic expressed in text. We use the term predictive sentiment analysis to denote the approach in which sentiment analysis is used to predict the changes in the phenomenon of interest. Our research goal is to investigate whether large-scale collections of daily posts from social networking and micro-blogging service Twitter are a suitable data source for predictive sentiment analysis. In our work we use the machine learning approach to learn a sentiment classifier for classification of financial Twitter posts (tweets) and causality analysis to show the correlation between sentiment in tweets and stock price movements. In addition, visual presentation of the sentiment time series for detection of important events is proposed. We analyzed financial tweets of eight companies (Apple, Amazon, Baidu, Cisco, Google, Microsoft, Netflix and Research In Motion Limited (RIM)) but due to space limitations, detailed analysis of only two companies (Google and Netflix) is presented in this paper.

The paper is structured as follows. Section 2 discusses Twitter specific text preprocessing options, and presents the developed Support Vector Machine (SVM) tweet sentiment classifier. The core of the paper is presented in Section 3 which presents the dataset collected for the purpose of this study, and the methodology developed for enabling financial market prediction from Twitter data. The developed approach proposes *positive sentiment probability* as an indicator for predictive sentiment analysis in finance. Moreover, by using the *Granger causality test* we show that sentiment polarity (positive and negative sentiment) can indicate stock price movements a few days in advance. Furthermore, since financial tweets do not necessarily express the sentiment, we have introduced sentiment classification using the *neutral zone*, which allows classification of a tweet into the neutral category, thus improving the predictive power of the sentiment classifier in certain situations. We conclude with a summary of results and plans for further work in Section 4.

2 Tweet Preprocessing and Classifier Training

In this work, we use a supervised machine learning approach to train a sentiment classifier, where classification refers to the process of categorizing a given observation (tweet) into one of the given categories or classes (positive or negative sentiment polarity of a tweet). The classifier is trained to classify new observations based on a set of class-labeled training instances (tweets), each described by a vector of features (terms, formed of one or several consecutive words) which have been pre-categorized manually or in some other presumably reliable way. This section describes the datasets, data preprocessing and the algorithm used in the development of the tweet sentiment classifier, trained from a set of adequately preprocessed tweets.

There is no large data collection available for sentiment analysis of Twitter data, nor a data collection of annotated financial tweets. For this reason, we have trained the tweet sentiment classifier on an available large collection of tweets annotated by positive and negative emoticons collected by Stanford University [18], approximating the actual positive and negative sentiment labels. This approach was introduced by Read [19]. The quality of the classifier was then evaluated on another set of actually manually labeled tweets.

To train the tweet sentiment classifier, we used a dataset of 1,600,000 (800,000 positive and 800,000 negative) tweets collected and prepared by Stanford University, where positive and negative emoticons serve as class labels. For example, if a tweet contains ":)", it is labeled as positive, and if it contains ":(", it is labeled as negative. Tweets containing both positive and negative emoticons were not taken into account. The list of positive emoticons used for labeling the training set includes :), :-), :), :D, and =), while the list of negative emoticons consists of :(, :-(, and : (. Inevitably this simplification results in partially correct or noisy labeling. The emoticons were stripped out of the training data for the classifier to learn from other features that describe the tweets. The tweets from this set do not focus on any particular domain.

The test data set collected and labeled by Stanford University contains tweets belonging to the different domains (companies, people, movies…). It consists of 498 manually labeled tweets, of which 182 were labeled as positive, 177 as negative and the others labeled as neutral. The tweets were manually labeled based on their sentiment, regardless of the presence of emoticons in the tweets.

As the Twitter community has created its own language to post messages, we explore the unique properties of this language to better define the feature space. The following tweet preprocessing options [18,20] were tested:

- **Usernames:** mentioning of other users by writing the "@" symbol and the username of the person addressed was replaced a unique token *USERNAME.*
- **Usage of Web Links:** web links were replaced with a unique token *URL.*
- **Letter Repetition:** repetitive letters with more than two occurrences in a word were replaced by a word with one occurrence of this letter, e.g., word *looooooove* was replaced by *love*.
- **Negations:** since we are not interested in particular negations, but in negation expressions in general, we replaced negation words (*not, isn't, aren't, wasn't, weren't,*

hasn't, haven't, hadn't, doesn't, don't, didn't) with a unique token *NEGATION*. This approach handles only explicit negation words and treats all negation words in the same way. Implicit negations and negative emotions presented in a tweet (e.g., A*void CompanyX*) are nevertheless handled to some extent by using unigrams and bigrams which assign negative sentiment to a word or a phrase in a tweet.

- **Exclamation and Question Marks:** exclamation marks were replaced by a token *EXCLAMATION* and question marks by a token *QUESTION*.

In addition to Twitter-specific text preprocessing, other standard preprocessing steps were performed [21] to define the feature space for tweet feature vector construction. These include text tokenization, removal of stopwords, stemming, N-gram construction (concatenating 1 to N stemmed words appearing consecutively) and using minimum word frequency for feature space reduction. In our experiments, we did not use a part of speech (POS) tagger, since it was indicated by Go et al. [18] and Pang et al. [22] that POS tags are not useful when using SVMs for sentiment analysis.

The resulting terms were used as features in the construction of TF-IDF feature vectors representing the documents (tweets). TF-IDF stands for term frequency-inverse document frequency feature weighting scheme [23] where weight reflects how important a word is to a document in a document collection.

There are three common approaches to sentiment classification [24]: (i) machine learning, (ii) lexicon-based methods and (iii) linguistic analysis. Instead of developing a Twitter-specific sentiment lexicon, we have decided to use a machine learning approach to learn a sentiment classifier from a set of class labeled examples. We used the linear Support Vector Machine (SVM) algorithm [25,26], which is standardly used in document classification. The SVM algorithm has several advantages, which are important for learning a sentiment classifier from a large Twitter data set: it is fairly robust to overfitting, it can handle large feature spaces [23,27] and it is memory efficient [28]. Given a set of labeled training examples, an SVM training algorithm builds a model which represents the examples as points in space separated with a hyperplane. The hyperplane is placed in such a way that examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a class based on which side of the hyperplane they are.

The experiments with different Twitter-specific preprocessing settings were performed to determine the best preprocessing options which were used in addition to the standard text preprocessing steps. The best classifier, according to the accuracy on the manually labeled test set, was obtained with the following setting: using N-grams of size 2, using words which appear at least two times in the corpus, replacing links with the *URL* token and by removing repeated letters in words. This tweet preprocessing setting resulted in feature construction of 1,254,163 features used for classifier training. Due to space limitations, the entire set of experimental results, including ten-fold cross-validation results, is not presented in the paper. Classifier testing showed that this preprocessing setting resulted in 81.06% accuracy on the test set, a result comparable to the one achieved by Go et al. [18].

3 Stock Market Analysis

This section investigates whether sentiment analysis on tweets provides predictive information about the values of stock closing prices. By applying the best classifier obtained with the process explained in Section 2, two sets of experiments are performed. In the first one, in which tweets are classified into two categories, positive or negative, the newly proposed sentiment indicators are calculated with the purpose of testing their correlation with the corresponding stock's closing price. We also present a data visualization approach used for detecting interesting events. In the second set of experiments the initial approach is advanced by taking into account the neutral zone, enabling us to identify neutral tweets (not expressing positive or negative sentiment) as those, which are "close enough" to the SVM's model hyperplane. This advancement improves the predictive power of the methodology in certain situations.

3.1 Data Used in the Stock Market Application

A large dataset was collected for these experiments. On the one hand, we collected 152,572 tweets discussing stock relevant information concerning eight companies in the period of nine months in 2011. On the other hand, we collected stock closing prices of these eight companies for the same time period. The data source for collecting financial Twitter posts is the Twitter API[2], i.e., the Twitter Search API, which returns tweets that match a specified query. By informal Twitter conventions, the dollar-sign notation is used for discussing stock symbols. For example, $GOOG tag indicates that the user discusses Google stocks. This convention simplified the retrieval of financial tweets. We analyzed English posts that discussed eight stocks (Apple, Amazon, Baidu, Cisco, Google, Microsoft, Netflix and RIM) in the period from March 11 to December 9, 2011. The stock closing prices of the selected companies for each day were obtained from the *Yahoo! Finance*[3] web site.

3.2 Sentiment and Stock Price Visualization

Using the best classifier obtained with the process explained in Section 2, we classified the tweets into one of two categories (positive or negative), counted the numbers of positive and negative tweets for each day of the time series, and plotted them together with their difference, the moving average of the difference (averaged over 5 days), and the daily stock closing price. The proposed visual presentation of the sentiment time series for Google can be seen in Fig. 1. Peaks show the days when people intensively tweeted about the stocks. Two outstanding positive peaks can be observed in August and October 2011. One is a consequence of Google buying Motorola Mobility and the other is due to recorded high increase in revenue and earnings for Google year-over-year.

[2] `https://dev.twitter.com/`
[3] `http://finance.yahoo.com/`

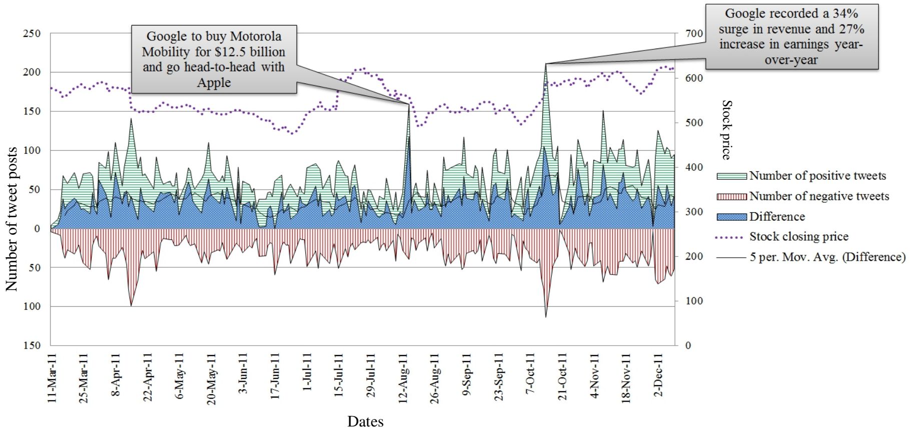

Fig. 1. Number of positive (green) and negative (red) tweet posts, their difference (blue), the moving average of the difference (averaged over 5 days), and the stock closing price per day for Google

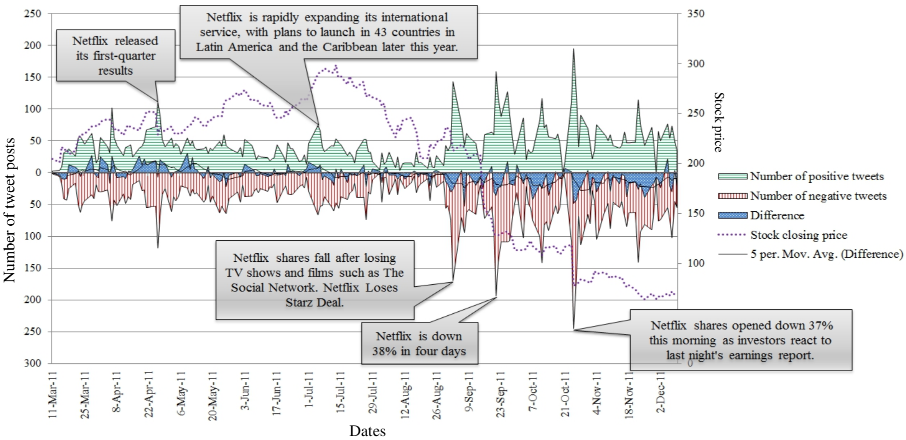

Fig. 2. Number of positive (green) and negative (red) tweet posts, their difference (blue), the moving average of the difference (averaged over 5 days), and the stock closing price per day for Netflix

This type of visualization can be used as a tool for easier and faster overview analysis of important events and general observation of trends. Relation of stock price and sentiment indicators time series provides insight into the reasons for changes in the stock price: whether they should be prescribed to internal market phenomena (e.g., large temporary buying/selling) or external phenomena (news, public events, social trends), which are expressed also in our sentiment indicators. It is interesting to observe how the tweet sentiment time series is often correlated with the stock closing price time series. For example, the sentiment for Netflix (Fig. 2) at the beginning of the year was mostly positive. As sentiment reversed its polarity in July, the stock closing price started to fall. For the whole second half of the year, sentiment remained mostly negative and the stock closing price continued to fall. On the other hand, a few days correlation between the sentiment and the stock closing price cannot be observed with the naked eye. For example, in Fig. 1 "Google buys Motorola" event has high amount of positive sentiments but stock price seems to drop. To calculate the real correlation we employ causality analysis, as explained in the next section.

3.3 Causality Analysis

We applied a statistical hypothesis test for stationary time series to determine whether tweet sentiment is related with stock closing price in the sense of containing predictive information about the values of the stock closing price or the other way around. To this end, we performed *Granger causality analysis* [29]. Time series X is said to Granger-cause Y if it can be shown that X values provide statistically significant information about future values of Y. Therefore, the lagged values of X will have a statistically significant correlation with Y. The output of the Granger causality test is the p-value. In statistical hypothesis testing, the p-value is a measure of how much evidence we have against the null hypothesis [30]; the null hypothesis is rejected when the p-value is less than the significance level, e.g., 5% ($p < 0.05$).

Positive Sentiment Probability. To enable in-depth analysis, we propose the positive sentiment probability sentiment indicator to be used in predictive sentiment analysis in finance. Positive sentiment probability is computed for every day of a time series by dividing the number of positive tweets by the number of all tweets on that day. This ratio is used to estimate the probability that the sentiment of a randomly selected tweet on a given day is positive.

Time Series Data Adaptation. To test whether one time series is useful in forecasting another, using the Granger causality test, we first calculated positive sentiment probability for each day and then calculated two ratios, which we have defined in collaboration with the Stuttgart Stock Exchange experts: (a) Daily change of the positive sentiment probability: positive sentiment probability today – positive sentiment probability yesterday, and (b) Daily return in stock closing price: (closing price today – closing price yesterday)/closing price yesterday.

Hypotheses Tested. We applied the Granger causality test in two directions, to test the following two null hypotheses: (a) "sentiment in tweets does not predict stock closing prices" (when rejected, meaning that the sentiment in tweets Granger-cause the values of stock closing prices), and (b) "stock closing prices do not predict sentiment in tweets" (when rejected, meaning that the values of stock closing prices Granger-cause the sentiment in tweets).

We performed tests on the entire 9 months' time period (from March 11 to December 9, 2011) as well as on individual three months periods (corresponding approximately to: March to May, June to August and September to November). In Granger causality testing we considered lagged values of time series for one, two and three days, respectively. The results indicate that in several settings sentiment of tweets can predict stock price movements. Results were especially strong for Netflix (Table 1), Baidu (all day lags for June-August and 2 and 3 day lags for March-May), Microsoft

(1 and 2 days lag for March-May and 1 day lag for the entire 9 months' time period), Amazon (2 and 3 days lag for September-November and 3 days lag for March-May) and RIM (all day lags for the entire 9 months' time period). For the given period, these companies had many variations in the closing price values (Baidu, Microsoft and Amazon) or a significant fall in the closing price (Netflix and RIM). On the other hand, the correlation is less clear for the other companies: Apple, Cisco and Google, which did not have many variations nor a significant fall in the closing price values for the given time period. This means that in the situations explained above, Twitter feeds are a suitable data source for predictive sentiment analysis and that daily changes in values of positive sentiment probability can predict a similar rise or fall of the closing price in advance.

Table 1. Statistical significance (*p*-values) of Granger causality correlation between daily changes of the positive sentiment probability and daily return of closing prices for Netflix

NETFLIX	Lag	Stocks = f(Tweets)	Tweets =f(Stocks)
9 months	1 day	0.1296	0.8784
March - May	1 day	0.9059	0.3149
June - August	1 day	0.1119	0.7833
September - November	1 day	0.4107	0.8040
9 months	2 days	**0.0067****	0.6814
March - May	2 days	0.4311	0.3666
June - August	2 days	0.2915	**0.0248****
September - November	2 days	**0.0007*****	0.9104
9 months	3 days	**0.0084****	0.6514
March - May	3 days	0.6842	0.3942
June - August	3 days	0.5981	**0.0734***
September - November	3 days	**0.0007*****	0.8464

*p < 0.1
**p < 0.05
***p < 0.001

Second Experimental Setup Results, Using the SVM Neutral Zone. In this section we address a three class problem of classifying tweets into the positive, negative and neutral category, given that not all tweets are either positive or negative. Since our training data does not contain any neutral tweets, we define a neutral tweet as a tweet that is "close enough" to the SVM model's hyperplane. Let us define the neutral zone to be the area along the SVM hyperplane, parameterized by t which defines its extent. Let d_{Pa} be the average distance of the positive training examples from the hyperplane and, similarly, let d_{Na} be the average distance of the negative training examples from the hyperplane. Then, the positive bound of the neutral zone is computed as

$$d_P(t) = t \cdot d_{Pa} \tag{2}$$

Similarly, the negative bound of the neutral zone is computed as

$$d_N(t) = t \cdot d_{Na} \tag{3}$$

If a tweet x is projected into this zone, i.e., $d_N(t)<d(x)<d_P(t)$, then it is assumed to bear no sentiment, i.e., that it is neutral. This definition of the neutral zone is simple and allows fast computation. Its drawback, however, is its lack of clear and general interpretation outside the context of a particular SVM classifier.

A series of experiments were conducted where the value for *t*, i.e., the size of the neutral zone, was varied and tested on manually labeled 182 positive and 177 negative tweets included in the test data set described in Section 2. Tweets were preprocessed using the best tweet preprocessing setting described in Section 2. With every new value of *t*, we calculated the accuracy on the test set and the number of opinionated (non–neutral) tweets, where the accuracy is calculated based on tweets that are classified as positive or negative by the SVM classifier described in Section 2. As a result of not taking neutral tweets into account when calculating the accuracy, the accuracy gets higher as we increase the *t* value (see Fig. 3). Hence, as we expand the neutral zone, the classifier is more confident in its decision about labeling opinionated tweets. As a negative side effect of increasing the neutral zone, the number of opinionated tweets is decreasing. We show this phenomena in Fig. 3, where also the number of opinionated tweets (classified as positive or negative) is plotted.

Accuracy is not a good indicator in the presented 3-class problem setting, therefore we experimentally evaluated the neutral zone according to its effect directly on stock price prediction. We repeated our experiments on classifying financial tweets, but now also taking into account the neutral zone. Since the Granger causality analysis showed that tweets could be used to predict movements of stock prices, we wanted to investigate whether the introduction of the neutral zone would further improve predictive capabilities of tweets. Therefore, every tweet which mentions a given company was classified into one of the three categories: positive, negative or neutral. Then, we applied the same processing of data as before (count the number of positive, negative and neutral tweets, calculate positive sentiment probability, calculate daily changes of the positive sentiment probability and daily return of stocks` closing price) and Granger analysis test. We varied the *t* value from 0 to 1 (where *t*=0 corresponds to classification without the neutral zone) and calculated average *p*-value for the separate day lags (1, 2 and 3). Results for Google and Netflix are shown in Fig. 4 and Fig. 5.

From Fig. 3 it follows that it is reasonable to focus on the parts of the plots that correspond to narrow boundaries of the neutral zone, for which the number of opinionated tweets is still considerable (for example up to 0.6). Slightly below 0.6 at 10% increase in accuracy, the loss of opinionated tweets is only at 20%. As for Google, shown in Fig. 4, the introduction of the neutral zone proved to be beneficial, as we got more significant *p*-values and the average *p*-value initially dropped. The best correlation between sentiment in tweets and stock closing price is observed at *t*=0.2 where

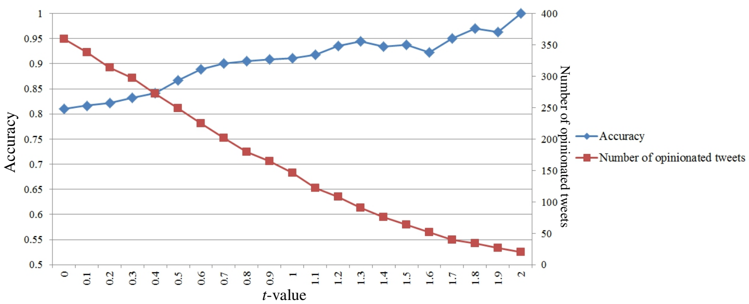

Fig. 3. The accuracy and the number of opinionated tweets while changing the value of *t*

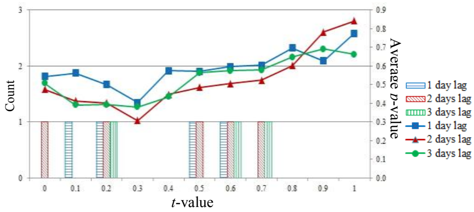

Fig. 4. Number of significant p-values (less than 0.1, column chart) and average p-values (line chart) for every day lag while changing the t value for Google

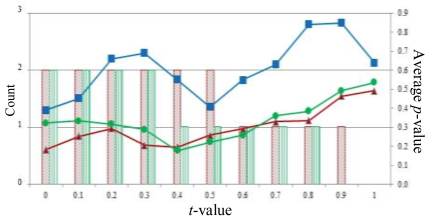

Fig. 5. Number of significant p-values (less than 0.1, column chart) and average p-values (line chart) for every day lag while changing the t value for Netflix

for every day lag we obtained a significant result and at t=0.3 where the average p-value is the lowest. Also for most of the other companies (Apple, Amazon, Baidu, Cisco and Microsoft), the neutral zone improved the predictive power of tweets, mostly for 2 and 3 days lags, given that the average p-value decreased when adding the neutral zone. Taking into account the number of significant p–values, the improvement was observed for Apple and Baidu, for Microsoft and Cisco there was a small and mixed improvement and for Amazon the number of p-values even dropped. In general, the best improvement was obtained with t=0.2.

For Netflix (see Fig. 5) and for RIM the neutral zone did not prove to be so useful. The neutral zone had the most positive effect on the results of Baidu and the least positive effects on the results for RIM. By analyzing these two extreme cases we observed that the closing price of Baidu had many variations and the RIM stock closing price has constant fall during a larger period of time. These observations of a relation between the impact of the neutral zone and the stock price variations hold also for most of the other companies.

In summary, it seems that when there is no apparent lasting trend of the closing price of a company, people write diverse tweets, with or without sentiment expressed in them. In such cases, it is desirable to use the neutral zone to detect neutral tweets in order to calculate the correlation only between the opinionated tweets and the stock closing price. On the other hand, it seems that once it is clear that the closing price of some company is constantly falling, people tend to write tweets in which they strongly express their opinion about this phenomenon. In this case, we might not need the neutral zone since there is no, or a very small number, of neutral tweets. The introduction of the neutral zone in such a situation may result in loss of information due to the tweets which get misclassified as neutral.

4 Conclusions

Predicting future values of stock prices is an interesting task, commonly connected to the analysis of public mood. Given that more and more personal opinions are made available online, various studies indicate that these kinds of analyses can be automated and can produce useful results. This paper investigates whether Twitter feeds are a suitable data source for predictive sentiment analysis. Financial tweets of eight companies (Apple, Amazon, Baidu, Cisco, Google, Microsoft, Netflix and RIM) were analyzed. The study indicates that changes in the values of positive sentiment probability can predict a similar movement in the stock closing price in situations where stock closing prices have many variations or a big fall. Furthermore, the introduced SVM neutral zone, which gave us the ability to classify tweets also into the neutral category, in certain situations proved to be useful for improving the correlation between the opinionated tweets and the stock closing price.

In future, we plan to experiment with different datasets for training and testing the classifier, preferably from a financial domain, in order for the classifier to be more finance adjusted since we are interested in this particular domain. Furthermore, we intend to expand the number of companies for further analysis to gain more insights in which situations our approach is most applicable. Finally, we plan to adjust our methodology to data streams with the goal to enable predicting future changes of stock prices in real-time.

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. The research was also supported by Ad Futura Programme of the Slovenian Human Resources and Scholarship Fund. We are grateful to Ulli Spankowski and Sebastian Schroff for their kind cooperation as financial experts in the stock analytics application presented in this paper.

References

1. Fama, E.: Random Walks in Stock Market Prices. Financial Analysts Journal 21(5), 55–59 (1965)
2. Kavussanos, M., Dockery, E.A.: Multivariate test for stock market efficiency: The case of ASE. Applied Financial Economics 11(5), 573–579 (2001)
3. Damasio, A.R.: Descartes error: emotion, reason, and the human brain. Harper Perennial (1995)
4. Nofsinger, J.R.: Social Mood and Financial Economics. Journal of Behavioral Finance 6(3), 144–160 (2005)
5. Gruhl, D., Guha, R., Kumar, R., Novak, J., Tomkins, A.: The predictive power of online chatter. In: Proceedings of the Eleventh ACM SIGKDD International Conference on Knowledge Discovery in Data Mining, pp. 78–87 (2005)
6. Tong, R.M.: An operational system for detecting and tracking opinions in on-line discussion. In: Working Notes of the ACM SIGIR 2001 Workshop on Operational Text Classification (OTC), pp. 1–6 (2001)
7. Mishne, G., Glance, N.: Predicting Movie Sales from Blogger Sentiment. In: AAAI Symposium on Computational Approaches to Analysing Weblogs AAAI-CAAW, pp. 155–158 (2006)
8. Asur, S., Huberman, B.A.: Predicting the Future with Social Media. In: Proceedings of the ACM International Conference on Web Intelligence, pp. 492–499 (2010)

9. Thelwall, M., Buckley, K., Paltoglou, G.: Sentiment in Twitter events. Journal of the American Society for Information Science and Technology 62(2), 406–418 (2011)
10. Ruiz, E.J., Hristidis, V., Castillo, C., Gionis, A., Jaimes, A.: Correlating financial time series with micro-blogging activity. In: Proceedings of the Fifth ACM International Conference on Web Search and Data Mining, pp. 513–522 (2012)
11. Bordino, I., Battiston, S., Caldarelli, G., Cristelli, M., Ukkonen, A., Weber, I.: Web search queries can predict stock market volumes. PLoS ONE 7(7), e40014 (2011)
12. Gilbert, E., Karahalios, K.: Widespread Worry and the Stock Market. In: Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media, pp. 58–65 (2010)
13. Bollen, J., Mao, H., Zeng, X.: Twitter mood predicts the stock market. Journal of Computational Science 2(1), 1–8 (2011)
14. Smailović, J., Grčar, M., Žnidaršič, M., Lavrač, N.: Sentiment analysis on tweets in a financial domain. In: 4th Jožef Stefan International Postgraduate School Students Conference, pp. 169–175 (2012)
15. Das, S., Chen, M.: Yahoo! for Amazon: Extracting market sentiment from stock message boards. In: Proceedings of the 8th the Asia Pacific Finance Association Annual Conference, APFA (2001)
16. Turney, P.: Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. In: Proceedings of the Association for Computational Linguistics, pp. 417–424 (2002)
17. Liu, B.: Sentiment Analysis and Opinion Mining. Morgan and Claypool Publishers (2012)
18. Go, A., Bhayani, R., Huang, L.: Twitter Sentiment Classification using Distant Supervision. In: CS224N Project Report, Stanford (2009)
19. Read, J.: Using emoticons to reduce dependency in machine learning techniques for sentiment classification. In: Proceedings of the ACL Student Research Workshop, pp. 43–48 (2005)
20. Agarwal, A., Xie, B., Vovsha, I., Rambow, O., Passonneau, R.: Sentiment analysis of twitter data. In: Proceedings of the Workshop on Languages in Social Media, pp. 30–38 (2011)
21. Feldman, R., Sanger, J.: The Text Mining Handbook - Advanced Approaches in Analyzing Unstructured Data. Cambridge University Press (2007)
22. Pang, B., Lee, L., Vaithyanathan, S.: Thumbs up?: sentiment classification using machine learning techniques. In: Proceedings of the ACL-2002 Conference on Empirical Methods in Natural Language Processing, vol. 10, pp. 79–86 (2002)
23. Joachims, T.: Text Categorization with Support Vector Machines: Learning with Many Relevant Features. In: Nédellec, C., Rouveirol, C. (eds.) ECML 1998. LNCS, vol. 1398, pp. 137–142. Springer, Heidelberg (1998)
24. Pang, B., Lee, L.: Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval 2(1-2), 1–135 (2008)
25. Vapnik, V.: The Nature of Statistical Learning Theory. Springer, New York (1995)
26. Cortes, C., Vapnik, V.N.: Support-Vector Networks. Machine Learning 20, 273–297 (1995)
27. Sebastiani, F.: Machine learning in automated text categoriztion. ACM Computing Surveys (CSUR) 34(1), 1–47 (2002)
28. Chang, C.C., Lin, C.J.: LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology (TIST) 2(3), 27 (2011)
29. Granger, C.W.J.: Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37, 424–438 (1969)
30. Schervish, M.J.: P Values: What They Are and What They Are Not. The American Statistician 50(3), 203–206 (1996)