STOCK MARKET PREDICTION USING SENTIMENT ANALYSIS ON TWITTER

Syamalekha CS, Sharath k, Sarath Sahadevan
B.Tech Computer Science department, Sahrdaya College of engineering & technology
Email - shyamalekhacs@rediffmail.com, sharathkunnath@gmail.com,sarathsahadev@outlook.com

Abstract: we apply sentiment analysis and machine learning principles to find the correlation between “public sentiment” and “market sentiment”. We use twitter data to predict public mood and use the predicted mood and previous days’ DJIA values to predict the stock market movements.

Key Words: sentiment analysis, twitter, facebook, stock market.

INTRODUCTION:

In this paper, we have a tendency to check a hypothesis supported the premise of behavioural political economy, that the emotions and moods of people aff ect their decison creating method, thus, resulting in an instantaneous correlation between ”public sentiment” and ”market sentiment”. We have a tendency to perform sentiment analysis on publically out there Twitter knowledge to find the general public mood and therefore the degree of membership into four categories - Calm, Happy, Alert and sort (somewhat like fuzzy membership). we have a tendency to use these moods and former days’ Dow-Jones Industrial Average Industrial Average (DJIA) values to predict future stock movements so use the anticipated values in our portfolio management strategy

LITERATURE REVIEW:

Our work is based on Bollen et al’s strategy which received widespread media coverage recently. They also attempted to predict the behavior of the stock market by measuring the mood of people on Twitter. The authors considered the tweet data of all twitter users in 2008 and used the OpinionFinder and Google Profile of Mood States (GPOMS) algorithm to classify public sentiment into 6 categories, namely, Calm, Alert, Sure, Vital, Kind and Happy. They cross validated the resulting mood time series by comparing its ability to detect the public’s response to the presidential elections and Thanksgiving day in 2008. They also used causality analysis to investigate the hypothesis that public mood states, as measured by the OpinionFinder and GPOMS mood time series, are predictive of changes in DJIA closing values. The authors used Self Organizing Fuzzy Neural Networks to predict DJIA values using previous values. Their results show a remarkable accuracy of nearly 87% in predicting the up and down changes in the closing values of Dow Jones Industrial Index (DJIA).

ALGORITHM:

The technique utilized in this paper builds directly on the one employed by Bollen. The raw DJIA values ar first fed into the preprocessor to get the processed values. At constant time, the tweets ar fed to the sentiment analysis algorithmic rule that outputs mood values for the four mood categories for every day. These moods and therefore the processed DJIA values ar then fed to our model learning framework that uses SOFNN to find out a model to predict future DJIA values victimisation them. The learnt model additionally because the previous DJIA and mood prices ar employed by the portfolio management system that runs the model to predict the long run value and uses the anticipated values to form acceptable buy/sell choices

SENTIMENT ANALYSIS:

1. Word List Generation We develop our own word list based on the well known Profile of Mood States (POMS) questionnaire. POMS is an established psychometric questionnaire which asks a person to rate his/her current mood by answering 65 diff erent questions on a scale of 1 to 5 (For example, rate on a scale of 1 to 5 how tensed you feel today?). These 65 words are then mapped on to 6 standard POMS moods- Tension, Depression, Anger, Vigour, Fatigue and Confusion. In order to do automate this analysis for tweets, the word list needs to be appropriately extended. Bollen et al. [1] used the Google n-grams data for the same. We followed a much simpler approach of extending the list by considering all commonly occuring synonyms of the base 65 words using SentiWordNet and a standard Thesaurus.

2. Tweet Filtering As mentioned earlier, the tweet data is enormous and will take several hours to be processed if used as it is (which makes the task of daily predictions difficult). Therefore, we filtered and considered only those tweets which are more likely to express a feeling, i.e. we consider only those tweets which contain the words ”feel”, ”makes me”, ”I’m” or ”I am” in them.

3. Score Mapping We map the score of each word to the six standard POMS states using the mapping techniques specified in the POMS questionnaire. We then map the POMS states to our four mood states using static correlation rules (for example, happy is taken as sum of vigour and negation of depression).

PORTFOLIO MANAGEMENT:

Having predicted the DJIA closing values one day in advance, we can use these predicted values to make intelligent sell/buy decisions. We develop a naive greedy strategy based on a simple assumption that we can hold at most one stock at any given time (or s stocks if all stocks are always bought and sold together) Following are the steps/features of our strategy

• Pre-computation we maintain a running average and standard deviation of actual adjusted stock values of previous k days

• Buy Decision If the predicted stock value for the next day is n standard deviations less than the mean, we buy the stock else we wait.

• Sell Decision If the predicted stock value is m standard deviations more than the actual adjusted value at buy time, we sell the stock else we hold.

CONCLUSION:

We have investigated the causative relation between public mood as measured from a large scale collection of tweets from twitter.com and the DJIA values. Our results show that firstly public mood can indeed be captured from the large-scale Twitter feeds by means of simple natural language processing techniques, as indicated by the responses

REFERENCES:

1. J. Bollen and H. Mao. Twitter mood as a stock market predictor. IEEE Computer, 44(10):91–94.
2. C.-C. Chang and C.-J. Lin. LIBSVM: A library for support vector machines. ACM Transactions on Intelligent Systems and Technology, 2:27:1–27:27, 2011.
3. G. P. Gang Leng and T. M. McGinnity. An on-line algorithm for creating self-organizing fuzzy neural networks. Neural Networks, 17(10):1477–1493.
4. Lapedes and R. Farber. Nonlinear signal processing using neural network: Prediction and system modeling. In Los Alamos National Lab Technical Report.
5. E. Stefano Baccianella and F. Sebastiani. Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In LREC. LREC.