Summer 6-26-2019

# Using Fine-grained Emotion Computing Model to Analyze the Interactions between Netizens' Sentiments and Stock Returns

Meng-xuan Zhu
*School of Information, Central University of Finance and Economics, Beijing, 100081, China*, 18222076336@163.com

Wei Zhang
*School of Information, Central University of Finance and Economics, Beijing, 100081, China*, kddzw@163.com

Meng Wang
*School of Information, Central University of Finance and Economics, Beijing, 100081, China*, financewm@163.com

Sui-xi Kong
*School of Information, Central University of Finance and Economics, Beijing, 100081, China*, ksx_cufe@163.com

Jian Wang
*School of Information, Central University of Finance and Economics, Beijing, 100081, China*, wanderingful@126.com

# Using Fine-grained Emotion Computing Model to Analyze the Interactions between Netizens' Sentiments and Stock Returns

*Meng-xuan Zhu, Wei Zhang*[*]*, Meng Wang, Sui-xi Kong, Jian Wang*
School of Information, Central University of Finance and Economics, Beijing, 100081, China

**Abstract:** From the perspective of behavioural finance, this paper combines the fine-grained sentiment calculation with the stock market econometric model to explore the interactions between netizens' sentiments and stock returns, analyze the differences in the influences of various emotions expressed by netizens on the stock market. First, it constructs a sentiment dictionary for the financial field; then, it calculates the emotion values contained in the text corpus, and constructs a textual sentiment classifier based on the recurrent neural network, calculates the emotion value and establishes the daily netizen sentiment index; and finally, it builds an econometric model to study the interactions between the netizen sentiment index and the stock returns. The results show that this model improves the accuracy of sentiment classification, reduces the number of iterations and saves computing resources; and that the netizen sentiment index, especially, "disgust" and "like", has significant effects on the stock price changes and transaction volumes, while on the other hand, the listed company's stock returns data has no reverse effect on the netizen sentiment index.

Keywords: Lexicon construction, Sentiment analysis, Emotion computing, Stock returns, Big data

## 1. INTRODUCTION

In the Web 2.0 era, social platforms such as Weibo and WeChat have become the main channels through which people express their opinions and exchange feelings. The short remarks on these platforms contain the comments and viewpoints of numerous Internet users, including their different sentiments towards the stock market, which show their personal mentalities, emotions, and subjective tendencies. To explore potential information containing emotional tendencies from these massive textual data has important theoretical and practical values for stock market forecasting[1], investment decision making[2]. Therefore, it has become a hot topic for researchers to explore the interactions between the netizens' sentiments and stock returns based on sentiment analysis.

At present, researches on the interactions between netizens' sentiments and stock returns all support the hypothesis that sentiment can affect the stock market. Based on a variety of sentiment proxy variables, the researchers divided the text sentiments into positive, negative and neutral ones, and analyzed the impacts of netizens' sentiments on stock market indices and abnormal rates of returns, etc. They believed that netizens' sentiments had significant effects on the stock returns, and that in turn, the stock returns also had reverse impacts on the netizens' sentiments, but how much is the reverse effect is still under theoretical discussion. Do netizens' sentiments affect the stock returns of a listed company? Does the stock returns have a reverse impact on netizens' sentiments? Are there any differences in the impacts of various sentiments on the stock returns? This series of questions remain to be studied.

Therefore, this paper attempts to collect the real stock data, measure the netizen sentiment index on the basis of the sentiment lexical ontology, construct a sentiment dictionary for the financial sector. We propose a new method to calculate emotion classification by using LTP platform. First, we classify the text corpus from the stock forum into fine-grained sentiment manually, and build a fine-grained sentiment classifier with the aid

[*] Corresponding author. E-mail: 18222076336@163.com (Meng-xuan Zhu), kddzw@163.com (Wei Zhang), financewm@163.com (Meng Wang), ksx_cufe@163.com (Sui-xi Kong), wanderingful@126.com (Jian Wang).

of long- and short-term memory network technologies, then compare the results of the two cases; and establish a statistical model containing the netizen sentiment index and stock market data to study the relationships between fine-grained sentiments and stock returns, so as to provide advice for the stock market investors .

## 2. RELATED WORK

### 2.1 The interaction between the netizens' sentiments and stock returns

Scholars studied the relationships between netizens' sentiments and stock price changes based on a variety of existing indices from different research perspectives, and found the impacts of netizens' sentiments on the stock returns, trading volume and non-systematic risks, etc. in the market[3].

Nardo et al. (2016) studied the predictive abilities of both Twitter and Google regarding the stock markets of various countries based on the discussion and search volumes and found that the Twitter indicator is ahead of the stock market returns in the U.S., U.K. and Canada while it is not quite predictive regarding the stock returns in Chinese market, but that there is a strong positive linear relationship between the Google indicator and the Chinese stock market[4]; Zhang and Yuan (2017) constructed an investor sentiment proxy indicator, which used the "Weibo index", and found the significant correlation between Weibo sentiments and the performance of the stock market through neural networks and other methods [5].Wang et al. (2017) believed there was a two-way causal relationship between netizens' sentiments and stock market returns and that the long-term impact was more significant [6]; Zhang et al. (2017), through the quantitative analysis of the fund comments in Sina Finance, proved that the investors' sentiments will affect the fund market yield, market volatility and trading volume [7]. Deng et al. (2017) studied the relationships between positive, negative and neutral sentiments and the Dow Jones Industrial Average, and found that negative sentiment has a certain impact on the hourly returns of stocks, but no impact on the daily returns [8].

From the existing research, it can be seen that both theory and practice support the hypothesis that sentiments affect the stock market. Researchers divided the textual sentiments into positive, negative and neutral ones based on a variety of different sentiment proxy variables, analyzed the impacts of netizens' sentiments on stock market indices and abnormal returns, etc., and believed that netizens' sentiments had significant effects on stock returns, and that stock returns also had a reverse impact on the changes in netizens' sentiments. However, such reverse impacts are still under theoretical discussion and lack empirical research.

### 2.2 The netizen's emotion value computing

The sentiment analysis of texts has been fully developed in the financial field and is mainly conducted through supervised learning and unsupervised learning. The traditional supervised classifier method is affected by many factors, and depends largely on a good corpus, and some algorithms require setting of a large number of parameters, which is not convenient for modelling. For this reason, scholars proposed an unsupervised method based on sentiment dictionary. For example, Jiang et al. (2015) built a social sentiment dictionary for Weibo based on the Weibo corpora contained in a multitude of social hot topics and compared the results of two different sentiment analysis methods to verify the effectiveness of the sentiment dictionary and the SVM sentiment analysis method[9]. Montejo-Ráez et al. (2014) proposed a sentiment analysis method combining sentiment calculation and random walk based on the graph structure of the dictionary WordNet. The experimental results showed that this method has better classification precision and recall than those of the traditional SVM method[10]. Dictionary-based methods are limited by dictionary coverage, and require large-scale corpora and heavy computation[11]. When dealing with large-scale short and unordered Weibo texts and extracting text features, they often face the problem of sparse features[12]; in addition, the existing sentiment classification focuses more on the three polarities (positive, negative and neutral) of sentiments, which are not fine-grained enough to fully characterize the overall evolution of Weibo sentiments [13].

## 3. RESEARCH DESIGN

**Step1. Data preparation**

We collect the data, including the reviews text corpus from the stock forum and stock returns data.

**Step2. Sentiment dictionary construction**

By referring to the method proposed by [9], this paper adopts the method based on the emotional lexical ontology and the deep learning algorithm Word2Vec. First, select the benchmark sentiment word so as to form a benchmark sentiment dictionary; then use Word2Vec to vectorize the words; adopt the incremental iterative process to achieve the extension of sentiment words; use the HowNet dictionary to filter out the words with high similarity, and then conduct manual screening, so as to obtain the benchmark sentiment dictionary.

**Step3. Calculation of the textual emotion value**

By reference to the method proposed by[14], use the "Language Technology Platform (LTP) " to calculate the emotion value of the text; use the corpus data as the input to the LTP platform; according to the result of the LTP dependency syntax analysis and financial sentiment dictionary, the emotion value is calculated according to different rules; then calculate the category of textual sentiment. For each piece of textual data, there will be scores of seven emotion categories.

**Step4. Classification of fine-grained sentiments**

- RNN classification: RNN (Recurrent Neural Network) is a kind of neural network with fixed weights, external input and internal state. In this paper, the LSTM network (Long- and Short-Term Memory network) is selected as a classifier of textual sentiments.
- ANN classification: In order to use emotion value as an input value for the classification, ANN (Artificial Neural Network) classifier is constructed. The ANN used in this experiment uses a fully connected layer to connect adjacent layers. The output of the output layers (including the hidden layer and the final output layer of the mode I) is shown in Formula (1).

$$o_j = \sum_{i=1}^{I} x_i \times w_{ij}, \ i = 1..100, j = 1..100 \tag{1}$$

In Formula (1), $o_j$ is the output of the model, $x_i$ is the input of the mode I, and $w_{i,j}$ is the weight between the ith node of the input layer and the jth node of the output layer. In this experiment, $x_i$ is a vector composed of the emotion value of certain text and the RNN output, and $o_j$ represents the corresponding emotion category of the text.

- Comparison of classification results: the classification results of the two classifiers are compared in terms of classification precision (iteration times and utilization rate of computing resources).

**Step5. Netizen sentiment index calcualtion**

The netizen sentiment index is divided into two categories – the weighted sum index of netizen emotion value and the average index of netizen emotion value, with a prefix of sum and avg, respectively. In order to better study the sentiments, the netizen sentiment index is subjected to three-day, five-day moving average processing, and 6 proxy variables of netizens' sentiments are obtained.

**Step6. Modelling**

Establish multiple models to determine the causal relationship between the netizen sentiment index constructed in this paper and the stock returns of the listed company, and analyze the specific form of impact of the netizen sentiment index on the stock returns of the listed company.

## 4. EXPERIMENTAL PROCESS AND RESULTS ANALYSIS

### 4.1 Experimental data

This paper selected four representative companies, namely FangDaTanSu, GuiZhouMaoTai, ZhongGuoPingAn, and KeDaXunFei (here in after referred to as their respective stock codes). The text corpus and stock returns of the four listed companies are obtained from the stock forum (www.eastmoney.com) and WIND database, respectively. The basic information of the listed companies and the text data volume of the stock forum are shown in Table 1.

**Table 1. Data of the listed companies**

| Company Name | GuiZhouMaoTai | FangDaTanSu | ZhongGuoPingAn | KeDaXunFei |
|---|---|---|---|---|
| Listing Date | 2001/8/27 | 2002/8/30 | 2007/3/1 | 2008/5/12 |
| Stock Code | 600519 | 600516 | 601318 | 2230 |
| Review start date | 2010/5/6 | 2010/5/7 | 2010/5/6 | 2014/11/6 |
| Total number of texts | 601506 | 683565 | 497083 | 364601 |
| Number of crawled texts | 321713 | 547211 | 165362 | 311846 |

Note: the statistics in this table are as of February 15, 2018.

For the four listed companies mentioned above, the data from October 8, 2015 to February 15, 2018 are captured, which contain a total of 1,346,132 pieces of textual data and 581 pieces of trading day data. For research needs, the sample data which contain textual data and trading day data are divided into two parts: training dataset, to construct the netizen sentiment index and the statistical model; and the test dataset, to test the validity and robustness of the index construction and the statistical model (shown in Table 2).

**Table 2. Different kinds of dataset used in the experiment**

| Data description | Start date | End date | Total days | Trading days | Amount |
|---|---|---|---|---|---|
| Total data | 2015/10/8 | 2018/2/15 | 862 | 581 | 1346132 |
| Training dataset | 2015/10/8 | 2017/5/31 | 602 | 402 | 312668 |
| Test dataset | 2017/6/1 | 2018/2/15 | 260 | 179 | 1033464 |

### 4.2 Experimental process

#### 4.2.1 Financial sentiment dictionary

With the help of the sentiment lexical ontology[15], the text is divided into seven emotion categories. The specific numbers of the emotion words are listed in Table 3. We can see that the number of words contained in each emotion category differs greatly after extension.

**Table 3. Number of words in each emotion category**

| Emotion category | Anger | Disgust | Fear | Happiness | Like | Sadness | Surprise | Total |
|---|---|---|---|---|---|---|---|---|
| Emotion lexicon | 388 | 10282 | 1179 | 1967 | 11108 | 2314 | 228 | 27466 |
| Ontology Corpus | 98 | 2867 | 319 | 637 | 2994 | 522 | 71 | 7508 |
| Benchmark dictionary | 90 | 2301 | 278 | 463 | 2572 | 529 | 53 | 6286 |
| Financial | 182 | 2636 | 354 | 489 | 2689 | 661 | 64 | 7075 |

For simplicity, by reference to the method adopted by [15,16], the emotion value in this paper is defined as the product of the sentiment intensity and the sentiment polarity, as shown in Formula (2):

$$emotion = intensity \times polarity \tag{2}$$

Where *intensity* means the emotion strength of the word, and its value is an integer from 1 to 9. *Polarity* is the polarity of word emotion. There are four sentiment polarities, namely 0 (neutral), 1 (positive), 2 (negative) and 3 (ambiguous). Emotion is the emotion value of a word, whose value is an integer from -9 to 9.

### 4.2.2 Netizen sentiment index

After the emotion value of each piece of text is obtained, the indices of the textual emotion value are collected every day, including the accumulated value, the average value, and the three-day and five-day moving average of the two. The symbols, calculation formulas and meanings are shown in Table 4.

**Table 4. Statistical indices of the emotion value**

| Name | Formula | Implication |
|---|---|---|
| *sum* | $sum_t = \sum_{i=0}^{I} sent_{i,t}$ | *Summary of each day emotion* |
| *sum_ma3* | $sum\_ma3_t = (\sum_{j=0}^{2} sumt_{t-j})/3$ | *3-day moving average of 'sum'* |
| *sum_ma5* | $sum\_ma5_t = (\sum_{j=0}^{4} sumt_{t-j})/5$ | *5-day moving average of 'sum'* |
| *avg* | $avg_t = (\sum_{i=0}^{I} sent_{i,t})/I$ | *Average of each day emotion* |
| *avg_ma3* | $avg\_ma3_t = (\sum_{j=0}^{2} avgt_{t-j})/3$ | *3-day moving average of 'avg'* |
| *avg_ma5* | $avg\_ma5_t = (\sum_{j=0}^{4} avgt_{t-j})/5$ | *5-day moving average of 'avg'* |

The correlation coefficients between the sentiment indices and the stock market index of KeDaXunFei (002230) show that the average emotion value indices and the stock market indices are not significantly correlated; On the other hand, the weighted sum indices of sentiment (sum, sum_ma3 and sum_ma5) and the market indices of the companies are significantly correlated. Therefore, this paper retains the weighted sum indices of sentiment as the sentiment indices for the subsequent establishment of the econometric model, and abandons the average indices of sentiment. For the dependent variable used in the following study, this paper selects the closing price (Cprice) and transaction volume (Volume) of the company as the proxy variables of the stock returns.

### 4.2.3 Establishment of the econometric model

- **Stationarity test**

Before modelling, the variables studied in this paper needs to be tested for stationarity. The method used in this paper is the unit root test – the ADF test. The current day's closing price (Cprice), transaction volume (Volume) and netizen sentiment index (sum_ma5) of the listed company undergo the ADF test and the test results are shown in Table 5. It can be found that Cpriceis not stationary at the significance level of 10%, while Volume and sum_ma5 remain stationary at the significance level of 1%. Therefore, the ADF test is performed on the first-order differential variable DCprice of the closing price Cprice, whose results are shown in Table 6. After the first-order difference, Cprice rejects the original hypothesis at the significance level of 1%, that is, the sequence is stationary. Therefore, when the current day's closing price is a single-order integrated sequence, the transaction volume and netizen sentiment index will be stationary sequences, and econometric models can be established for the variables DCprice, Volume and sum_ma5 to verify their influence relationships (as shown in

Table 7 ).

**Table 5. ADF test results of the original variables**

| Variable | t-value | p-value | Critical Value 1% | Critical Value 5% | Critical Value 10% | Conclu. |
|---|---|---|---|---|---|---|
| *Cprice* | -2.425 | 0.1355 | | | | Non-stationary |
| *Volume* | -4.603*** | 0.0002 | -3.449 | -2.87 | -2.571 | Stationary |
| *sum_ma5* | -4.021*** | 0.0015 | | | | Stationary |

Note: in the table, *** indicates p<0.01; ** indicates p<0.05; and * indicates p<0.5

**Table 6. First-order differential ADF test results**

| Variable | t-value | p-value | Critical Value 1% | Critical Value 5% | Critical Value 10% | Conclu. |
|---|---|---|---|---|---|---|
| *DCprice* | -19.905*** | 0 | -3.449 | -2.87 | -2.571 | Stationary |

Note: in the table, *** indicates p<0.01; ** indicates p<0.05; and * indicates p<0.5

- **Granger causality test**

The Granger causality test is used to study the relationship between the netizen sentiment index and the stock returns of the listed company. The results of the Granger causality test on the variables of 002230 are shown in Table 7. The row variables in Table 7 are the causes of the Granger causality, and the column variables are effects of the Granger causality. For example, the original hypothesis in the first row and the second column is: DCprice is not the Granger cause of Volume.

**Table 7. Granger causality test results**

| p-value | DCprice | Volume | sum_ma5 |
|---|---|---|---|
| *DCprice* | - | 0.7403 | 0.6607 |
| *Volume* | 0.0406 | - | 0.3721 |
| *sum_ma5* | 0.0941 | 0.0119 | - |

According to Table 7 at the significance level of 0.1, sum_ma5 is the Granger cause of DCprice; at the significance level of 0.05, sum_ma5 is the Granger cause of Volume, and Volume is the Granger cause of DCprice. The rest of Granger causality test results are not significant.

- **Simultaneous Formulas model**

From the results of stationarity test and Granger causality test, it can be found that two Formulas need to be established: Formula(3) describes the impacts of Volume and sum_ma5 on DCprice, and Formula(4) describes the impact of sum_ma5 on Volume.

$$DCprice_t = \alpha_0 + \alpha_1 Volume_t + \alpha_2 sum\_ma5_t + \varepsilon_t \tag{3}$$

$$Volume_t = \beta_0 + \beta_1 sum\_ma5_t + \mu_t \tag{4}$$

Then, regression is performed on the experimental data, and the resulting model is shown as the econometric model I in Table 8. It can be seen that $\alpha_2$ in Formula(3) is significant at the significance level of 0.1, and that the other constant terms and coefficients are significant at the significance level of 0.01. It is not difficult to see that the coefficient $\alpha_1$ in Formula(3) is too small; in other words, Volume has little impact on DCprice. So now we remove Volume from the Formula(3) and build a new simultaneous Formulas model. The result is shown as the econometric model II in Table 8. The coefficient of $\alpha_1$ becomes insignificant, indicating that Volume should not be removed from the model.

**Table 8. Simultaneous Formulas model results**

| Coeff. | Model I | Model II |
|---|---|---|
| $\alpha_0$ | 27.936*** | 30.789*** |
| $\alpha_1$ | 0.0001*** | - |
| $\alpha_2$ | -0.00496* | 0.004802 |
| $\beta_1$ | 22609.13*** | 22609.13*** |
| $\beta_2$ | 77.37366*** | 77.37366*** |

Note: in the table, *** indicates p<0.01; ** indicates p<0.05; and * indicates p<0.5

Regarding the experimental data set, the goodness of fit $R^2$ of Formula(3) in the econometric model I is 0.320, and the adjusted $R^2$ is 0.316; the goodness of fit $R^2$ in the econometric model II is 0.006, and the adjusted $R^2$ is 0.004. In both models, $R^2$ of Formula(4) is 0.078 and the adjusted $R^2$ is 0.076.

- **Fine-grained emotion model**

According to the results of the Granger causality test on various sentiment indices and stock returns data, ang_ma5, dis_ma5 and lik_ma5 are the Granger causes of DCprice, while hap_ma5 is not Granger causes of DCprice, nor is DCprice the Granger cause of each sentiment index. Therefore, an econometric model is established for the DCprice and sentiment indices ang_ma5, dis_ma5 and lik_ma5, as shown in Formula(5). The influence coefficient of ang_ma5, dis_ma5 is significant at the confidence level of 5%, while that of lik_ma5 is significant at the confidence level of 1%.

$$DCprice=38.405-0.0382ang\_ma5-0.062dis\_ma5+0.035lik\_ma5 \qquad (5)$$

$$(1.144) \quad (0.020)^{*} \qquad (0.036)^{*} \qquad (0.014)^{**}$$

The goodness of fit $R^2$ of the above model is 0.556 and the adjusted $R^2$ is 0.548. As can be seen, the sentiment indices ang_ma5, dis_ma5 and lik_ma5 have significant effects on the stock price.

### 4.3 Experiment analysis

According to the correlation analysis of the sentiment indices and the stock returns indices, the correlations between the weighted sum indices of sentiment and the stock returns are significant at different levels, while the correlations between the average indices of sentiment and the stock returns are not significant, mainly due to the following two reasons: on the one hand, the weighted sum indices of sentiment takes all textual emotion values into account - the better the stock returns is, the more sentiments the netizens will express; on the other hand, regarding the average indices of sentiment, the denominator is the total amount of all the text published on the day, of which a large portion does not contain sentiments. Such text is only noise to the average indices.

In the selection of the sentiment proxy indices, after screening, it is found that the correlation coefficients between the weighted sum of emotion values, three-day moving average of the weighted sum of emotion values and five-day moving average of the weighted sum of emotion values and the current day’s closing price of the company are 0.790, 0.839 and 0.875, respectively, and that the correlation coefficients with respect to the transaction volume are 0.765, 0.736 and 0.746, respectively. It can be seen that the five-day moving average of the weighted sum of emotion values well fits the current day’s closing price and transaction volume of the company, as the moving average of emotion values reduces the interferences caused by short-term fluctuations, For this reason, this paper selects the five-day moving average of the weighted sum of emotion values as the proxy variable of the netizen sentiment index.

On the basis of correlation analysis, Granger causality test and econometric analysis are performed on the relationship between the netizen sentiment index and the stock returns. The Granger causality test results show that the transaction volume and the netizen sentiment index have significant impacts on the current day’s closing price, and the netizen sentiment index also has a significant impact on the transaction volume. The coefficient of the Volume is 0.0001, when we remove Volume from the Formula(3) and build a new model, The coefficient of $\alpha_1$ becomes insignificant, indicating that Volume should not be removed from the model. The coefficient of the netizen sentiment index is -0.00496, which is significant at the confidence level of 10%.

The correlations between the fine-grained emotion values and the stock returns are significant. Except that the amount of price increase/decrease, the correlation coefficients of other indices are all significant at the confidence of 0.05, with the maximum one being up to 0.903. The Granger causality test results show that only dis_ma5 and lik_ma5 in the fine-grained sentiment variables are the Granger causes of DCprice, while DCprice is not the Granger cause of any fine-grained sentiment index. The econometric model for the DCprice and the sentiment indices dis_ma5 and lik_ma5 shows that the coefficients are both significant at the confidence level of 1%, 0.1127 for dis_ma5 and 0.1318 for lik_ma5, which means that the expression of disgust will have a negative effect on the stock price, while "like" will help increase the stock price.

## 5. CONCLUSIONS

Based on the large amount of stock text data, we build a sentiment dictionary for the financial field, calculate emotion values, and apply it in the fine-grained sentiment classification. Then an econometric model is proposed to explore the relationship between the netizens' sentiment and the stock returns. The results show that: (1) by using the deep learning-based fine-grained sentiment classifier, the model has higher classification precision and reduced iteration times, which saves computing resources; (2) the weighted sum indices of emotion values are significantly correlated with the stock returns, while the average indices of emotion values are not significantly correlated with the stock returns; (3) the netizen sentiment index has significant effects on the stock price changes and transaction volumes of the listed company, while on the other hand, the listed company's stock returns has no reverse effect on the netizen sentiment index; (4) each fine-grained sentiment index has a significantly different impact on the stock price. The emotion, "disgust" and "like", have significant impacts on stock price changes, while others do not, nor do stock price changes affect the fine-grained sentiment.

Based on the results of this study, future research should focus on further improving the financial sentiment dictionary, building the multi-grain sentiment classifier and optimizing classifier parameters.

## ACKNOWLEDGEMENT

This work was supported by the National Natural Science Foundation of China (71874215,71571191), Beijing Natural Science Foundation (9182016,9194031) and MOE (Ministry of Education in China) Project of Humanities and Social Sciences (15YJCZH081,17YJAZH120,19YJCZH253).

## REFERENCES

[1] Sul H K, Dennis A R, Yuan L I. (2016). Trading on twitter: Using social media sentiment to predict stock returns. Decision Sciences, 48(3): 454-488

[2] Xu X, Yang Z, Hao L. (2017).Research on the risk evaluation of the risk investment project based on the investor sentiment with triangular fuzzy information. Journal of Intelligent & Fuzzy Systems, 33(6): 3201-3208

[3] Tetlock P C, Saar‐Tsechansky M, Macskassy S. (2008). More than words: Quantifying language to measure firms' fundamentals. The Journal of Finance, 63(3): 1437-1467

[4] Nardo M, Petracco‐Giudici, M Naltsidis. (2016). Walking down wall street with a tablet: A survey of stock market predictions using the web. Journal of Economic Surveys, 30(2): 356-369

[5] Xindong Z , Dongliang Y. (2017). Research on the Impact of Investor Sentiment on Stock Market Based on Micro-blog. Journal of Intelligence, 36(8): 81-87

[6] Wang F L, Wang X Y. (2017).Does social emotion affect stock return—The evidence from sina weibo. Journal of Shanxi Finance and Economics University, 39(2): 35-46

[7] Zhang W Y, Meng N D, Zhang T P. (2017).Research on the Stock Market Volatility Influenced by Institutional Investor in China: Based on Investor Sentiment. Journal of Beijing University of Technology (Social Sciences Edition), 17(1):60-66

[8] Deng S, Huang Z, Sinha A P, et al. (2017).The Interaction between Microblog Sentiment and Stock Return: An Empirical Examination. MIS quarterly, 42(3): 895-918

[9] Jiang S Y, Huang W J, Cai M L, Wang L X. (2015).Building Social Emotional Lexicons for Emotional Analysis on Microblog. Journal of Chinese Information Processing,29(6): 166-171

[10] Montejo-Ráez, Arturo, Martínez-Cámara, Eugenio, Martín-Valdivia, M Teresa, et al. (2014). Ranked wordnet graph for sentiment polarity classification in twitter. Computer Speech & Language, 28(1): 93-107

[11] Huang F L, Feng S, Wang D L, Yu G. (2017).Mining Topic Sentiment in Microblogging Based on Multi-feature Fusion. Chinese Journal of Computers,40(4): 872-88

[12] Sun X, Peng X Q, Hu M, Ren F J. (2017).Extended Multi-modality Features and Deep Learning Based Microblog Short Text Sentiment Analysis. Journal of Electronics & Information Technology,39(9): 2048-55

[13] Li X, Li J, Wu Y. (2015).A global optimization approach to multi-polarity sentiment analysis. Plos One,10(4):e0124672

[14] Che W X, Li Z H, Liu T. (2010).Ltp: A Chinese Language Technology Platform. In: Proceedings of the Coling 2010: Demonstrations, Beijing: Tsinghua University Press, 13-16

[15] Xu L H, Lin H F, Zhao J. (2008).Construction and Analysis of Emotional Corpus. Journal of Chinese Information Processing, 22(1): 116-122

[16] Zhang W, Kong S X, Zhu Y C, et al. (2018).Sentiment classification and computing for online reviews by a hybrid SVM and LSA based approach. Cluster Computing, 2018, online in press.