

# Research on the Transfer Rules of Internet Users' Negative Emotional State in Financial Public Opinion

Lei Li

The Third Construction Co., Ltd. of China Construction Eighth Engineering Division, Nanjing, China  
Email: 3369717@qq.com

**How to cite this paper:** Li, L. (2020) Research on the Transfer Rules of Internet Users' Negative Emotional State in Financial Public Opinion. *Open Journal of Business and Management*, 8, 282-301.  
<https://doi.org/10.4236/ojbm.2020.81017>

**Received:** December 2, 2019

**Accepted:** January 3, 2020

**Published:** January 6, 2020

Copyright © 2020 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

As the rapid development of internet and the booming of financial market in China, the study of extracting the emotional state of netizens from financial public opinions and using it for quantitative investment analysis has drawn a lot of attention. Because of the limitation of datasets scale, quantitative investment analysis based on financial public opinion has some unsolved problems in the research of financial analysis, such as the results cannot predict the stock price in real stock markets. Based on the long-short-term memory network in deep learning, the proposes study combined with the theory of herding effect in behavioral finance, this paper designs an emotional classification model for netizens' comments on social media, interpret emotional state transaction of netizens through sentiment analysis, forming an investor's emotional states' transfer model, and incorporating the emotional states as a factor into the stock price-forecasting model at last. The results show that the investor's emotional states have a significant impact on stock price volatility. This stock price forecasting method based on sentiment analysis also provides a new technical path for quantitative investment analysis in the financial market.

## Keywords

Financial Public Opinion, Netizen, Negative Emotional State, Transfer Rule

---

## 1. Introduction

Internet technology and its application are gradually penetrating into our lives. The continuous innovation of Internet model, the acceleration of online and offline service integration and the acceleration of online public service are all

promoting the continuous growth of Internet users. The Internet has significant convenience and interactivity. With the rise and development of various information, open and professional forums, people began to express various opinions on the Internet, resulting in a large number of public opinion information (Du Hongtao 2016) [1].

With the gradual improvement of national cultural level, the popularization of stock market policy and the rapid rise of various media, more and more people begin to use the Internet for investment and financing, and express their views on the Internet, which integrate financial public opinion. Behavioral finance research shows that emotions have a significant impact on investors' investment decisions (Zhou 2018; Liu Weiwei and Liu Xinxin 2014) [2] [3]. In a positive mood, investors tend to overestimate investment opportunities and underestimate investment risks, making transactions frequent; in a negative mood, on the contrary. Such irrational investment behavior is not conducive to the stable development of the stock market. Therefore, how to correctly analyze financial public opinion information and use it to correctly guide investors' investment sentiment is particularly important.

This paper creatively combines the herd effect theory in behavioral finance, designs an emotional classification model for netizens' comments in social media, and obtains the emotional state index through emotional analysis. At the same time, this paper explains the emotional state transfer of Internet users, forms the emotional state transfer model of investors, and discusses the relationship between emotional state of Internet users and stock price in a certain period of time. A stock market forecasting model based on netizens' emotions is established and has achieved good performance, which provides reference support for investors' investment decisions.

## 2. Literature Review

### 2.1. Behavioral Finance Theory

Since modern financial theory is no longer sufficient to explain various investment phenomena in the stock market, behavioral finance theory is gradually popularized and has become a research hotspot in various countries. Behavioral finance holds that investors are limited rational. The existence of emotional and cognitive biases makes them unable to achieve rational expectation and utility maximization. Their limited rational behavior will lead to market inefficiency and asset price deviation from basic value (Asano and Taeibanaki 1992) [4]. At present, the research content of behavioral finance is mainly divided into three parts: individual behavior of investors (Lan *et al.* 2018) [5]; group behavior of investors (Xiang Cheng and Lu Jing 2018) [6]; limited arbitrage and ineffective market. At present, the research of behavioral finance is only from the perspective of market performance and psychology, lacking of objective quantitative research. This study will analyze the influence of netizens' investment sentiment on the stock market and the transfer rules of netizens' emotional state from the

perspective of data and the “herd effect” theory [7] in behavioral finance.

## 2.2. Emotion Classification of Financial Public Opinions

The study of financial public opinion first needs to classify the emotions of netizens in the financial field. The function of the emotion classification is to extract the text based on an specialized event, theme, object, thing and other object content from the text information, extract the author’s subjective evaluation emotion and attitude, and judge the emotional orientation of the text (Kennedy and Inkpen 2010; Mamei *et al.* 2016) [8] [9]. According to the classification methods, emotion classification can be divided into five categories: dictionary based methods (Zhang *et al.* 2018) [10], machine learning based methods (Abdul *et al.* 2018) [11], hybrid methods of dictionary and machine learning, methods based on weak annotation information and methods based on deep learning (Sohangir *et al.* 2018) [12].

Deep learning is a method based on representation learning of data in machine learning. Existing research shows that deep learning is suitable for public opinion analysis (Xi Xuefeng and Chen Guodong 2016; Chen Weihua and Xu Guoxiang 2018; SuZhi *et al.* 2017) [13] [14] [15]. Based on the deep learning method, this paper will use the long short term memory (LSTM), which performs well in time series prediction and text classification, to classify financial public opinion texts. Word2vec is used as word vector construction tool. Word embeddings method which proposed by Mikolov *et al.* (Mikolov *et al.* 2013) [16] have been widely used in Chinese word segmentation, emotional classification, syntactic dependency analysis, etc. (Socher *et al.* 2013) [17] [18].

## 2.3. Research on Emotional Application

Investor sentiment refers to irrational noise traders’ false belief in the unpredictability of future returns on risky assets. Sarlin (2017) [19] analyzed the bank crisis based on news data, semantic analysis and positive and negative discrimination of news text using neural network. Chang and Wang (2018) [20] used the emotional analysis method to explore the degree of risk in the sharing economy. Krishnamoorthy (2018) [21] used the concept of financial and non-financial performance indicators to propose a hierarchical sentiment classifier model based on association rules, which uses performance indicators to analyze the sentiment of financial news articles. Tian Gaoliang *et al.* (2018) [22] empirically tested the relationship between online public opinion and the information efficiency of listed companies. These emotional messages mainly come from two sources, which used to come mainly from the news [23] and now mainly from the growing social media.

Investors, as netizens, express their opinions and emotions by Posting on the Internet. Such public opinion information often reflects investors’ views and expectations on the stock market (Brown and Cliff 2004) [24]. A large number of behavioral finance studies and experiments show that investor sentiment has a significant impact on investment decisions when there are uncer-

tainties and risks in investment (Li *et al.*, 2014; Wen *et al.*, 2018; Hsu and Liao, 2016; Xu Yuandong, 2015) [25] [26] [27] [28]. Shi Shanchong *et al.*, (2018) [29] studied the relationship between investor sentiment time series, closing price and trading volume of Shanghai stock index based on wechat text mining, and found that the proportion of negative emotions can stably predict the closing price of Shanghai stock index, and the proportion of positive emotions and neutral emotions of investors. The increase or decrease of trading volume of Shanghai Stock Exchange index can quickly lead to the increase or decrease of trading volume of Shanghai Stock Exchange Index lagging for one day. Lai Kaisheng *et al.* (2014) [30] modified the emotion financial decision-making model, which well explains the role of investor sentiment in financial decision-making. Therefore, it is theoretically feasible to quantify investor sentiment and predict the stock market with this feature.

Gilbert *et al.* [31] estimated anxiety, worry and fear from 20 million LiveJournal posts, and found that the increase in these negative expressions could predict downward pressure on the s & p index. Oh *et al.* [32] used content on Stock-Twits and Yahoo Finance to predict stock prices. They build sentiment classifiers based on artificially labeled bullish, bearish and neutral emotions, and use sentiment indices to predict the future direction of the stock market.

Of course, perfect financial forecasts do not exist, and the actual stock market is often influenced by many factors. Tsibouris [33] pointed out in the paper that when the accuracy of prediction is higher than 56%, it is considered to have achieved a relatively ideal effect, which has a good guiding role for future stock investment. In this experiment, LSTM model is used to creatively add investor sentiment index on the basis of original variables to predict stock price trend and verify the effect.

## 3. Method

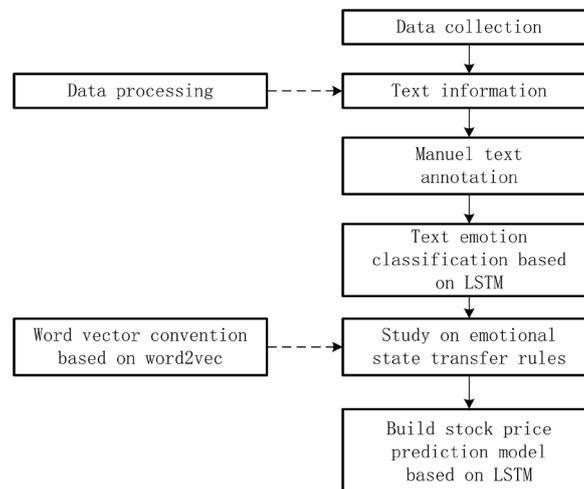
### 3.1. Model Design

#### 3.1.1. Framework Design

In this paper, we will conduct Emotional Analysis on the financial public opinion text, analyze the transfer rules of netizens' negative emotional state in the financial public opinion, and finally give a stock price prediction model based on Netizens' emotional state. The overall design framework is shown in **Figure 1**.

The model design process is as follows:

- 1) Using data grabbing system to collect the data of "Guba".
- 2) Data structure and noise reduction processing.
- 3) According to the marking rules, the text is manually marked with positive and negative emotions.
- 4) Word2vec is used to segment the text, construct the emotion classification model, and carry out the emotion classification experiment.
- 5) Study the rules of negative emotional state transition.
- 6) Forecasting stock price based on LSTM and emotional state.



**Figure 1.** Research design.

### 3.1.2. Technical Framework

With the upsurge of deep learning research, various deep learning frameworks emerge in endlessly. Good frameworks can improve learning efficiency and quality. According to GitHub, the most popular deep learning framework is TensorFlow, followed by Keras and Caffe.

TensorFlow, as a relatively high-level machine learning library, users can easily use it to design neural network structure, without having to write C++ or CUDA code in order to achieve high efficiency. Like Theano, it supports automatic derivation, and users do not need to solve gradients through back propagation. TensorFlow supports not only the common Network structures: Convolutional Neural Network (CNN) and Recurent Neural Network (RNN), but also deep reinforcement learning and other computation-intensive scientific calculations (such as partial differential equation solution).

Keras is more of a minimalist, highly modular neural network library. Keras aims to enable users to conduct the fastest prototype experiments and minimize the process of turning ideas into results. Theano and Tensorflow's calculations support more general-purpose calculations, while Keras specializes in deep learning. The biggest problem of keras is that it cannot directly use multi GPU at present, so the large-scale data processing speed is not as fast as other frameworks supporting multi GPU and distributed.

In this paper, TensorFlow is selected as the back-end, and all algorithm programs are supported by TensorFlow. At the same time, Keras, which is relatively simple and modular, is chosen as the front-end platform, and TensorFlow is used as the experimental environment of this study.

### 3.1.3. Data Annotation Rules

This paper classifies the public opinion data into three categories: positive emotion, negative emotion and neutral emotion. Due to the special emotional characteristics of different fields of corpus, especially the emotional words in the Internet financial industry, there are many proper nouns that can reflect the

emotional tendency, so the traditional tagging corpus is not fully suitable for this study. In view of the above situation, the following emotional tagging rules are formulated for the Internet financial industry. All texts are labeled according to the following rules in **Table 1**.

## 3.2. Data Collection and Processing

### 3.2.1. Data Selection

In traditional public opinion analysis (Shi Shanchong 2018; Wu Qinglin and Zhou Tianhong 2016) [29] [34], social media platforms such as microblog, forum and wechat are often used as data sources, however, this study focuses on the relationship between netizens' emotional tendency and the stock market in the field of Finance and economics. Therefore, it is necessary to include professional financial forum into the consideration of data acquisition, and has high requirements for data quality and the normalization and authority of data platform.

At present, the main financial websites in China include eastmoney, sina finance, hexun, tonghuashun, China economic and zhongjinzaixian. Through the analysis of comprehensive ranking, Baidu weight and content of financial websites, combined with the characteristics of data professionalism, reliability and real-time, this paper chooses eastmoney and Sina financial as data source websites.

In the face of a large number of listed companies, this study selected LETV (300104), Guizhou Maotai (600519) and Zhangzidao (002069) as research samples. This paper mainly studies the influence of netizens' emotion on the stock price in the stock market, so it is necessary to obtain the comment text that can reflect netizens' emotion and the stock price information of specific stocks in the opening time.

### 3.2.2. Information Collection

According to the particularity of the research industry, this research develops a set of information collection system and embeds the information grabbing function into the financial public opinion analysis system, which is developed by Vue + node + mysql. In view of the uniform presentation rules of the required public opinion texts and the fact that the information platforms are all web sites, the node language, which has a strong ability to parse the web pages, is selected as the development language of the information collection system.

**Table 1.** Emotion labeling rules.

classification	Classification rules
Positive	Text with obvious bullish mentality, positive tone optimistic, with "bullish", "blue chip", "positive" and other positive words.
Negative	Text with obvious bearish mentality, negative tone, more swear words, with "fall", "diving", "bad" and other negative words.
neutral	Text has no obvious emotional color, simple statement tone or text has nothing to do with the stock market.

### 3.2.3. Data Preprocessing

Using the above self-developed information grabbing system to grab data, taking LETV as an example, a total of 314,344 pieces of data were finally obtained, including 62,255 pieces of Sina stock bar data and 252,089 pieces of eastmoney stock bar data. All data is stored in mysql database. The information fields retrieved include the time of the comment, the link of the comment, the title of the post, the amount of reading, the body of the post and the author. Because the acquired text information contains a lot of noise, it is necessary to preprocess the data. The processing rules are as follows:

#### 1) Structured processing

All data are sorted by time series according to the purpose of analysis. At the same time, the repetitive text is filtered, that is, if the same text appears more than once, the text with the earliest occurrence time is retained and the remaining duplicate text is deleted.

#### 2) Content processing

The content processing of data is mainly reflected in three aspects. First of all, the public opinion information acquired in the text, a large number of the “official” text, this text often contains a lot of words, and presentation content has nothing to do with the stock to analyze, through the analysis of a large number of texts, found the comment text of individual shareholders is usually short, therefore in the process of data processing to delete characters more than 100 public opinion data. Second, there is no lack of text containing special symbols and null values in the collected data. In order to obtain more useful information and prevent noise interference, special symbols and empty text records need to be deleted. Thirdly, the research results of this paper are applicable to all stocks, which requires that all texts should not contain the vocabulary specific to one stock when processing data. Therefore, in the process of processing, this study will delete the relevant stock abbreviation, code and other proprietary words in the public opinion text to prevent errors in the emotional analysis.

## 4. Empirical Analysis

### 4.1. Sentiment Classification of Stock Market Reviews Based on LSTM

#### 4.1.1. Text input Representation

In this paper, 25,111 texts were selected from the obtained public opinion information texts for annotation, among which 6400, 9363 and 9348 were labeled as positive, negative and neutral texts respectively, and their corresponding emotional states were labeled as 1, -1, 0, and stored in the excel table.

The word segmentation tool used is the jieba word segmentation, and the dictionary used is the dictionary of the jieba word segmentation. Due to the particularity of the financial industry, many proprietary words cannot be identified, which will affect the efficiency of word segmentation. To solve this problem, this paper constructs a dictionary of financial proprietary words based on the financial industry. It has been proved by practice that the efficiency of word segmen-

tation can be significantly improved and the result of word segmentation basically meets the expected effect. Then, word2vec is used for vector transformation of the words

#### 4.1.2. Experimental Results

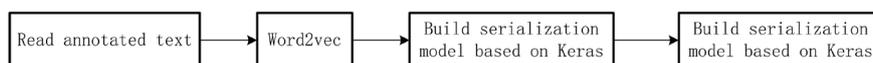
The algorithm was written according to the process shown in **Figure 2**. Firstly, annotated text was read, word segmentation was carried out by word2vec, Keras serialization model was established, and various parameters of LSTM model were finally set.

After a period of training, the model training was finished. The training results are shown in **Figure 3**.

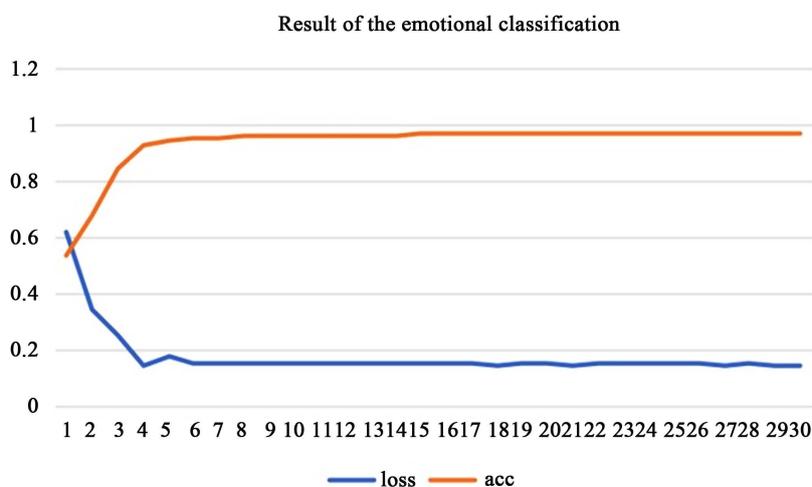
It can be seen from **Figure 3** that after 7 rounds of training, the model is basically stable, the accuracy of the training set is about 97%, and the loss rate is stable around 0.14. In the test set, the accuracy and loss rate was 95.12% and 0.16%, respectively. In general, the model has certain practicability.

To prevent over fitting is a big difficulty in machine learning. It means that in a statistical model, too many feature dimensions lead to a perfect training set of fitting functions, but the prediction results of new data are poor. In this experiment, the size of Dropout is adjusted to prevent the model from over fitting. The value of Dropout is to make a neuron stop working as a probability, so as to improve the generalization ability of the network, so as to achieve the purpose of preventing over fitting and improving the prediction accuracy. The test results are shown in **Table 2**.

It can be seen from **Table 2**, when Dropout < 0.5, the training accuracy of the model decreases, while the test accuracy increases. When Dropout = 0.6, the test accuracy decreases. The above phenomenon indicates that when Dropout = 0.5,



**Figure 2.** Algorithm flow chart.



**Figure 3.** Training results.

**Table 2.** Experimental results with different dropout.

Dropout	train accuracy	test accuracy
0.1	99.21%	91.95%
0.2	98.87%	92.29%
0.3	98.66%	92.84%
0.4	98.44%	94.39%
0.5	97.75%	95.12%
0.6	96.98%	94.86%

the effect of overfitting prevention is the best. Therefore, 0.5 was used as Dropout parameter in this experiment to prevent over-fitting.

The trained model weights and network structure were saved in the model file, and the remaining 300,000 unmarked texts were emotionally annotated. After automatic annotation, the annotated texts were stored in the mysql database according to time series for the following research on the emotional relationship between stock price and netizens.

## 4.2. Research on Emotional State Transfer

This paper makes a preliminary hypothesis on the relationship between investor's investment sentiment and stock price based on the theory of "herd effect" and emotion-financial decision model in behavioral finance, and holds that investor's sentiment can affect stock price, and in this process, the negative mood state transfer rules of Internet users are affected by "herd effect". Next, we do empirical research on this hypothesis.

### 4.2.1. Data Processing

Take "LETV", "Guizhou Maotai" and "Zhangzi Dao" as examples to collect trading data in recent years, including opening price, closing price, maximum and minimum price, and trading volume. Add the proportion of negative emotions in the public opinion of LETV on that day to form the basic data. The data format is shown in **Table 3**.

### 4.2.2. Stock Market Prediction Model Experiment Based on Emotional Analysis

In this experiment, the closing price, trading volume and the proportion of negative emotions in the first ten days are used as input  $x$ , and the closing price in the next day is used as output  $y$  for in-depth learning based on time series. The selected model is LSTM. The following is a description of the parameters of the main steps in the experiment:

```

model = Sequential() //Initialize model
model.add(LSTM(6, input_shape = (10,2))) //Set time step to 10, feature to 2
model.add(Dense(output_dim = 1)) //The output dimension is 1
model.add(Activation('linear')) //The activation function is "linear"

```

The training results of "LETV" are shown in **Figure 4**.

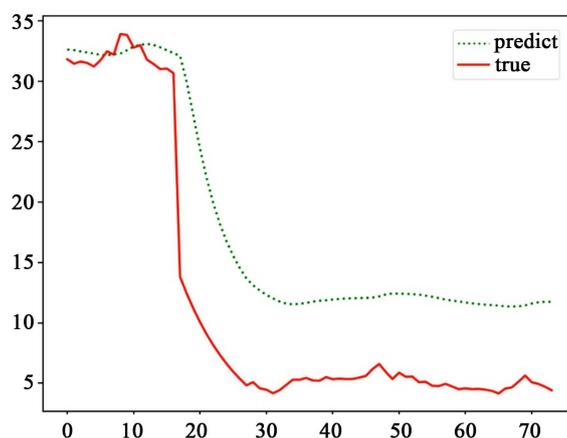


Figure 4. Training results of the stock market prediction model “LETV”.

Table 3. Prediction model data format.

Time	Maximum price	Minimum price	Opening price	Closing price	Trading volume	Negative sentiment ratio
2015/11/05	49.88	53.29	51.49	50.37	852701	2.175
2015/11/06	50.03	55.02	50.03	53.16	866227	1.8889
2015/11/09	51.41	57.39	52.5	55	852201	2.1667
...	...	...	...	...	...	...
2018/04/27	4.37	4.76	4.37	4.59	2121350	1.9345

The loss rate of the model is close to 0, and the basic trend of the stock price is basically in line with the expectation, but the stock price is not accurate enough. Especially after LETV experienced the “capital chain” storm, LETV has been crumbling, investors have become extremely cautious, “herd effect” makes buyers and sellers become irrational, so many factors make the stock cannot be predicted in accordance with the general method.

Take this stock as an example to verify the validity of investor sentiment. By adjusting various parameters of the model, the prediction accuracy of stock price is closer to the real value after adding the characteristic of investors’ investment sentiment. Figure 5 compares the trend of stock price prediction before and after adding the proportional feature of negative emotions.

It can be seen from the above comparison that after adding the characteristic variable of emotion, the predicted stock price is more accurate. In order to verify that the above model is not an accident in an ideal state, this study forecasts the stock price of “Guizhou Maotai” and “Zhangzi Dao”, and the prediction results are shown in Figure 6 and Figure 7.

From the perspective of data, the accuracy of forecasting the rise or fall reached 59.11%, the accuracy of predicting the rise or fall within 1% reached 54.14%, and the accuracy of predicting the rise or fall within 2% reached 73.52%. Table 4 shows the statistical table of accuracy before and after the application of this index.

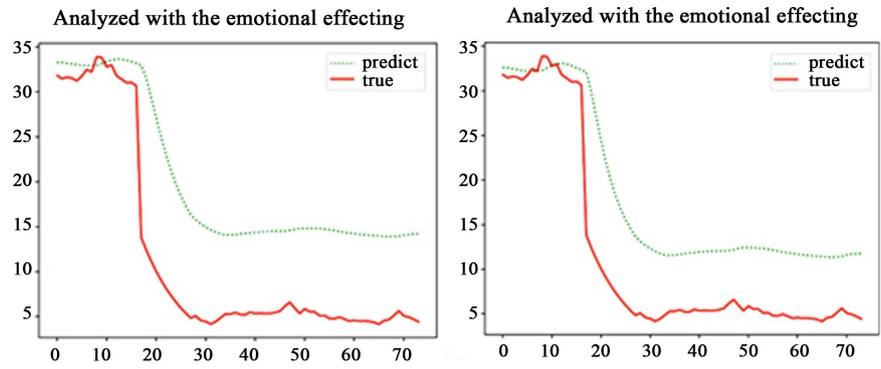


Figure 5. Prediction comparison chart.

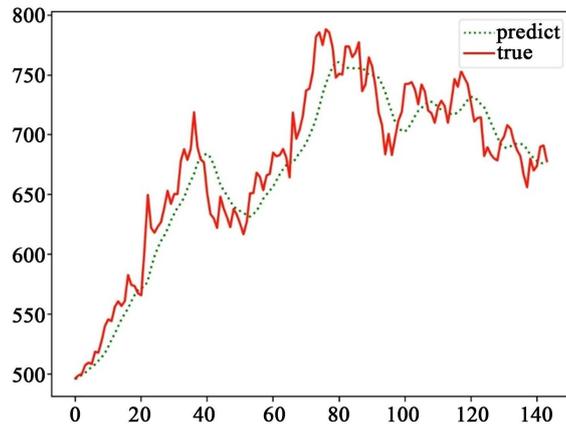


Figure 6. Prediction of "Guizhou Maotai".

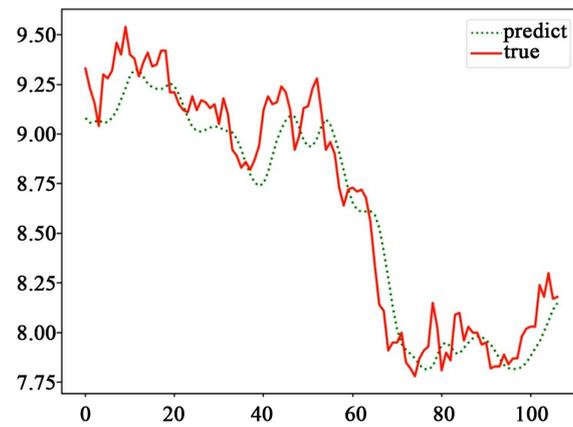


Figure 7. Prediction of "Zhangzi Dao".

Table 4. Forecast statistics.

Statistical indicators	Accuracy	
	Without emotional indicators	With emotional indicators
Up/down prediction	58.25%	59.11%
within 1%	52.14%	54.14%
within 2%	68.66%	73.52%

### 4.2.3. Stock Market Prediction Model Experiment Based on Emotional Analysis

In order to better prove the above conclusions, this paper selected panel model to conduct root variance test on the experimental results. The panel model can validate each stock individually. In this experiment, variables were selected as the opening price, closing price, maximum price and minimum price in the fundamentals, as well as the proportion of public opinion attention and negative emotions in the public opinion surface, and multiple logistic regression algorithm was selected for regression analysis. The algorithm is as follows:

$$p_i = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \dots + \beta_6 * X_6 + \varepsilon \quad (1)$$

$X_1$  to  $X_6$  represent the opening price, the closing price, the highest price, the lowest price, the proportion of public opinions and negative emotion ratio, respectively, and  $\beta_0$  to  $\beta_6$  are the weight of each indicator variable every day. The research object and situation of this paper, the above model can be extended into the following two models:

Variable weight verification model:

$$p_i = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \dots + \beta_6 * X_6 + \varepsilon \quad (2)$$

Time weight validation model:

$$p_a = \alpha_0 + \alpha_1 * p_{i1} + \alpha_2 * p_{i2} + \dots + \alpha_{10} * p_{i10} + \gamma \quad (3)$$

The variable weight model is mainly used for regression analysis of various parameters in the stock price prediction model to observe the influence of various indexes on stock price fluctuation.  $X$  represents the value of each indicator (such as closing price, opening price, etc.), and  $\beta$  is the weight of each indicator variable every day. Time weighting model is mainly used for regression analysis of time weighting, where  $p_a$  represents the actual value of the predicted stock price and  $\alpha$  represents the day. In the above stock price prediction model, the first 10 days are selected as the prediction range in this study, so it is necessary to verify the impact of time on the accuracy of the model.

In order to find important influence factors and prove the reliability of the model, this study embedded the variable weight verification model into the time weight verification model, and verified the correlation between various indicators and the predicted results through multiple logistic regression. The experimental results are shown in **Table 5** and **Table 6**.

As can be seen from the above table, in terms of time dimension, it generally conforms to the rule that the shorter the time distance is, the more important the information is. In terms of influence factor dimension, among the four important trading indicators, the maximum price has the greatest influence on the predicted results. The influence weights of public opinion attention and negative emotion ratio on the predicted results reach 0.4222 and 0.3211 respectively, which have significant influence on the test results and are not negligible. Therefore, public opinion attention and negative emotion ratio can affect the stock price.

**Table 5.** Time weight analysis.

Time parameter	Max price
$a_1$	0.5415
$a_2$	0.3519
$a_3$	0.4410
$a_4$	0.3911
$a_5$	0.1622
$a_6$	0.1941
$a_7$	0.4512
$a_8$	0.1345
$a_9$	0.3291
$a_{10}$	0.2311

Note:  $a_1 - a_{10}$  represents the first 1 - 10 days of the forecast.

**Table 6.** Variables weight analysis.

Time parameter	Max price
Opening price	0.4315
Closing price	0.4925
Highest price	0.6171
Minimum price	0.5611
Public opinion attention	0.4222
proportions of negative emotions	0.3211

#### 4.2.4. Emotional State Transition Model

In the above paper, the influence of investor sentiment on stock price was verified. In order to verify the influence of Internet users' sentiment on stock price, policies and investment sentiment of people around them, the following emotional state transfer model hypothesis was proposed:

$$T_0 = \beta_1 * T_1 + \beta_2 * T_2 + \beta_3 * N_1 + \beta_4 * N_2 + \beta_5 * PR_1 + \beta_6 * PR_2 + \beta_7 * P_t + \varepsilon \quad (4)$$

where  $T_0$  represents the proportion of negative emotions on the day, and  $T_1$  and  $T_2$  represent the proportion of negative emotions on the day before and two days before, respectively.  $PR_1$  and  $PR_2$  represent the rise and fall of stock prices in the previous day and two days respectively.  $N_1$  and  $N_2$  represent the attention of the previous day and two days respectively,  $P_t$  represents whether there is a major policy. The specific labeling rules are shown in **Table 7**.

In order to study whether there is correlation between each variable and the dependent variable, correlation analysis is needed. The analysis results are shown in the following **Table 8**.

As can be seen from the above table, apart from the relatively low correlation between policies and stock prices due to individual reasons, the remaining several have obvious resonance. Regression analysis is conducted on the above variables.

**Table 7.** Explanation of tagging rules.

parameter	explanation of parameter
Negative emotion ratio (T)	Divide the negative emotion scale into 5 scales, When $T < 1$ , $T = 1$ When $1 < T < 2$ , $T = 2$ When $2 < T < 3$ , $T = 3$ When $3 < T < 4$ , $T = 4$ When $T > 4$ , $T = 5$
Attention of Internet users (N)	divide netizens into three types According to the number of posts, When $N < 300$ , $N = 1$ When $300 < N < 700$ , $N = 2$ When $N > 700$ , $N = 3$
Range of price rise and fall (PR)	The range of the stock price rise or fall, When $PR < -2.5\%$ , $PR = 1$ When $-2.5\% < PR < 0$ , $PR = 2$ When $0 < PR < 2.5\%$ , $PR = 3$ When $PR > 2.5\%$ , $PR = 4$
Policy (PI)	Major favorable policies, $PI = 1$ Major adverse policies, $PI = -1$ No major policies, $PI = 0$

**Table 8.** Descriptive statistics of variables.

	$T_1$	$T_2$	$PR_1$	$PR_2$	$N_1$	$N_2$	$P_1$
T1	1						
T2	0.821	1					
PR1	0.662	0.572	1				
PR2	0.463	0.701	0.252	1			
N1	0.141	0.044	0.210	0.117	1		
N2	0.071	0.651	0.154	0.176	0.286	1	
PI	0.820	0.713	0.651	0.351	0.108	0.056	1

Perform logistic multivariate logistic regression according to the data format previously processed, and the regression equation is as follows:

$$T_0 = 0.612 + 0.422 * T_1 + 0.334 * T_2 + 0.071 * N_1 + 0.012 * N_2 - 1.211 * PR_1 - 0.927 * PR_2 - 0.104 * P_1 \quad (5)$$

The negative emotions were negatively correlated with the rise and fall, that is, the larger the rise, the lower the negative emotions. Negative emotions are positively correlated with the proportion of negative emotions in the previous days. Emotions affect each other in the investor group, and herd behavior is obvious. Although there is a positive correlation between negative emotions and the number of posts, the correlation coefficient is small and the number of posts has a limited effect on the judgment and prediction of negative emotions. Therefore, this variable is deleted from the model. Negative sentiment is negatively correlated with policies, that is, favorable policies can guide investor sentiment to a

positive direction. The adjusted model is as follows:

$$T_0 = 0.612 + 0.422 * T_1 + 0.334 * T_2 - 1.211 * PR_1 - 0.927 * PR_2 - 0.104 * P_l \quad (6)$$

#### 4.2.5. Interpretation of Negative Emotional State Transition Rules

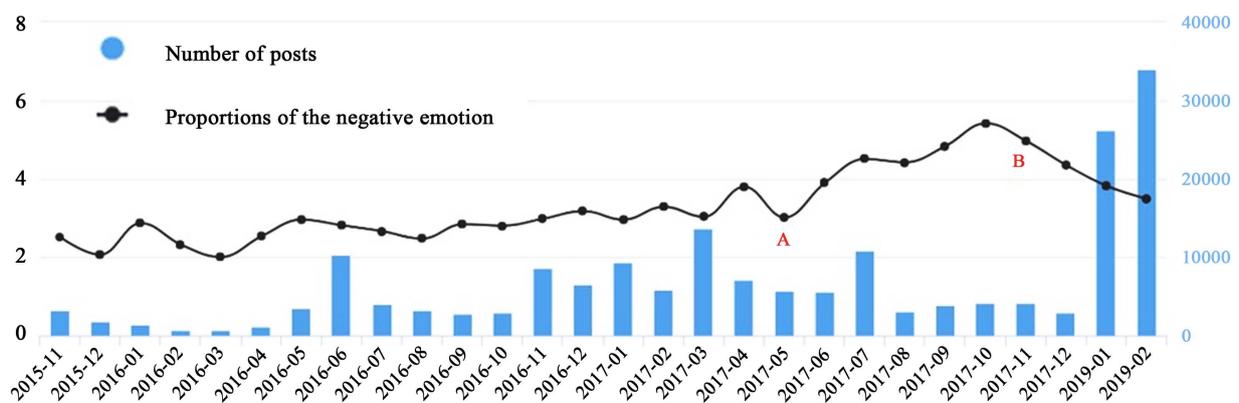
In order to explain the above model, this paper takes “LETV” as an example to conduct emotional statistics on the marked data text. The proportion of positive and negative emotions and the corresponding number of texts are shown in **Table 9**, among which the proportion of negative emotions is as high as 45%, the proportion of neutral emotions is 42%, and the proportion of positive emotions is only 13%. It is not difficult to find that in the posts of the stock bar, the text with negative emotions is far more than the text with positive emotions.

In the investment world, sentiment is a concept of uncertainty, which reflects the willingness or expectation of market participants to invest. When the earnings do not meet expectations, that is, investors are pessimistic about market information (capital, policy) will produce negative emotions. On the other hand, when investors have doubts about the prospects of the company and feel not optimistic about the operating conditions and the prospects of the company, they will also have negative emotions. The logic of investor sentiment’s influence on future market fluctuations lies in the amplification of positive and negative news step by step, and the joint efforts of various factors (bad news and good news) will transfer the current investor sentiment state to the next stage.

This experiment takes “Letv” as an example and takes month as the unit to analyze “Letv” from two perspectives of Posting quantity and emotional state. Due to the small number of posts before November 2015, they were not included in the statistical range. The statistical results are shown in **Figure 8**.

**Table 9.** Text emotional state distribution.

Classification	Text number	Emotional proportion
Positive emotions	24,976	13%
Neutral emotions	79,009	42%
Negative emotions	84,414	45%



**Figure 8.** Statistics of netizens’ emotions and the number of posts.

From the perspective of quantity, it is not difficult to see that the number of netizens' posts declined sharply after March 2017, but suddenly showed a blowout increase in January 2018. The reason is that "Letv" suspended its trading on April 14, 2017, and resumed trading successfully on January 31, 2018. Blank in the last eight months of relatively quiet period netizens, and after the resumption of Letv became the one big heat that the stock market community discussion, suspended during the netizen emotions relative depression, all sorts of bad news from the company continues to be strengthened, and the good news is weakening, leads to analyze shares have drop stop, netizens have in stock right to express their opinions and feelings. Similarly, during December 2015 to June 2016, "Letv" also experienced suspension of trading, and the stock price tendency also appeared relatively silent phenomenon.

From the perspective of emotional state, the negative emotions of netizens as a whole have been on the rise. Meanwhile, Letv's share price continues to fall. In **Figure 8**, A significant emotional turning point is May 2017 (point A) and November 2017 (point B). Point A is that "Letv" has just experienced A suspension of the wave of company scandals. Investors could not sell their shares, they were disappointed with Letv, and the proportion of negative emotions began to rise sharply. After point B, the emotional ratio of netizens began to decline, because the news of "Letv" resumption of trading began to spread. After letv resumed trading on January 31, 2018, its share price continued to drop by the daily limit. In this period of time, most people are irrational, negative emotions to strengthen, blindly sell the stock, lack of their own a long-term rational judgment.

It can be seen from **Figure 8** that the emotional state of Internet users is closely related to the company's operating conditions, prospects and other factors. During the period when the proportion of netizens' negative emotions rose, Letv's stock price also gradually declined. The ratio of negative emotions and stock prices showed a negative correlation. Netizen mood is low reflect in reduce hold status, sell off wait for a behavior to go up. When most people are pessimistic about the prospect of Letv, the original opinions of Letv supporters begin to waver, which is embodied in the rising proportion of negative emotions and the continuous decline of the stock limit. This also well explain the "flock effect" in the behavioral finance theory, in this kind of extremely negative investment environment, due to the lack of the understanding of information and doubts about their point of view, investors are difficult to uncertainty make reasonable expectations for the future of the market, but by observing the behavior of other investors around and view to extract information, in this kind of information continuously, many information will be roughly the same and reinforce each other, resulting in a herd behavior.

Through the experiment and the comparison between the stock price and the transfer curve of investors' emotional state, it can be found that the netizens' investment sentiment can affect the stock price, but it can also be affected by the stock price and other investors' sentiment.

## 5. Discussion

In the financial field, investors' investment sentiment has a significant impact on the stock market, and the level of investors' investment sentiment often influences the order of the whole financial market. According to the characteristics of the financial field, this paper constructed a financial dictionary for the Internet financial industry, developed a crawler system for "stock bar" and other financial and economic BBS, and realized the efficient capture of all financial text data. Based on deep learning, this study classifies investors' emotional posts on financial BBS such as "stock bar" to obtain the emotional index of relevant stocks, and proposes an emotional state transfer model on this basis. In this paper, while making reasonable explanations for the transfer of emotional state, this emotional index was incorporated into the stock prediction model, and the conclusion was drawn as follows:

1) Sentiment analysis can affect stock prices.

In the existing studies on the influence of emotions on stock prices (Zengrunxi *et al.* 2014; Zheng Xiaoxue and Chen Fuji, 2015) [35] [36], the selection of variables and the annotation of data are usually highly subjective, so the constructed emotional indicators lack certain scientific procedures and basis. In this paper, LSTM was used to build a stock prediction model based on the transaction indicators and emotional ratio of letv, Kweichow Moutai and Zhangzi Dao. Model is based on ten days before the closing price, ten days before the opening price, the highest and the lowest price and the emotional state to predict the next day's share price, with no join emotional indicators, comparing the model prediction results found in the price of the stock prediction and volatility has obviously improved on prediction accuracy, regression analysis is carried out on the time dimension and the impact factor, again to verify the mood which also affects prices. Compared with traditional model, this model has small deviation and high stability.

2) The transfer of investors' emotional state is affected by many factors at the same time.

Traditional emotional state transition model usually quantifies a number of indicators manually and calculates them by means of artificial calculation. Such quantification often contains subjective factors, which are likely to cause errors and thus affect the trend of state transition. This paper innovatively combined the "herd effect" theory in behavioral finance to argue that people are not all rational, that is, people are easily influenced by others when making investment decisions. Therefore, this experiment incorporated the emotional influence of people around into the model as a variable factor. After the regression analysis of stock price rise and fall, investor investment sentiment, public attention and policy, the author selected rise and fall, investor investment sentiment and policy as model variables and finally constructed the emotional state model as shown in formula 6. The experimental results show that the model is practical and has a significant improvement in interpretability and stability compared with the traditional model.

The stock forecasting method based on emotion analysis also provides a new way for quantitative investment in financial market.

## Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

## References

- [1] Du, H.-T., Meng, Q.-G. and Wang, J.-Z. (2016) Effectiveness of Internet Data in the Public Opinion Analysis Task. *China Soft Science*, No. 4, 34-44.
- [2] Zhou, Z.K., Xu, K. and Zhao, J.C. (2018) Tales of Emotion and Stock in China: Volatility, Causality and Prediction. *World Wide Web-Internet and Web Information Systems*, **21**, 1093-1116. <https://doi.org/10.1007/s11280-017-0495-4>
- [3] Liu, W.-Q. and Liu, X.-X. (2014) Individual/Institutional Investor Sentiment and Stock Returns: Study Based on Shanghai A-Share Market. *Journal of Management Sciences in China*, **17**, 70-87.
- [4] Asano, S. and Tabei-naki, T. (1992) Testing the Constancy of Relative Risk Aversion: An Analysis of Japanese Household Financial Asset Data. *Journal of Japanese and International Economies*, **6**, 52-70. [https://doi.org/10.1016/0889-1583\(92\)90018-Y](https://doi.org/10.1016/0889-1583(92)90018-Y)
- [5] Lan, Q.J., Xiong, Q.Y., He, L.J. and Ma, C.Q. (2018) Individual Investment Decision Behaviors Based on Demographic Characteristics: Case from China. *PLoS ONE*, **13**, e0201916. <https://doi.org/10.1371/journal.pone.0201916>
- [6] Xiang, C. and Lu, J. (2018) Investor Limited Attention, Industrial Information Diffusion and Stock Pricing. *Systems Engineering-Theory & Practice*, **38**, 817-835.
- [7] Ma, L. (2016) An Empirical Test of The Herding Effect: Evidence from the China Stock Market. *Nankai Economic Studies*, No. 1, 144-153.
- [8] Kennedy, A. and Inkpen, D. (2010) Sentiment Classification of Movie Reviews. Using Contextual Valence Shifters. *Computational Intelligence*, **22**, 110-125. <https://doi.org/10.1111/j.1467-8640.2006.00277.x>
- [9] Ma, M., Liu, D.-S. and Li, H. (2016) Research on the Network Public Opinion Analysis System Model Based on Big Data. *Information Science*, **34**, 25-28+33.
- [10] Zhang, S.X., Wei, Z.L., Wang, Y. and Liao, T. (2018) Sentiment Analysis of Chinese Micro-Blog Text Based on Extended Sentiment Dictionary. *Future Generation Computer Systems*, **81**, 395-403. <https://doi.org/10.1016/j.future.2017.09.048>
- [11] Abdul, A., Chen, J.H., Liao, H.Y. and Chang, S.H. (2018) An Emotion-Aware Personalized Music Recommendation System Using a Convolutional Neural Networks Approach. *Applied Sciences*, **8**, 1103. <https://doi.org/10.3390/app8071103>
- [12] Sohangir, S., Wang, D.D., Pomeranets, A. and Khoshgoftaar, T.M. (2018) Big Data: Deep Learning for Financial Sentiment Analysis. *Journal of Big Data*, **5**, 3-28. <https://doi.org/10.1186/s40537-017-0111-6>
- [13] Xi, X.-F. and Zhou, G.-D. (2016) A Survey on Deep Learning for Natural Language Processing. *Acta Automatica Sinica*, **42**, 1445-1465.
- [14] Chen, W.-H. and Xu, G.-X. (2018) Research on the Prediction Accuracy of Stock Market Volatility Based on Deep Learning and Stock Forum Data. *Management World*, **34**, 180-181.
- [15] Su, Z., Lu, M. and Li, D.-X. (2017) Deep Learning in Financial Empirical Applica-

- tions: Dynamics, Contributions and Prospects. *Journal of Financial Research*, No. 5, 111-126.
- [16] Mikolov, T., Chen, K., Corrado, G., *et al.* (2013) Efficient Estimation of Word Representations in Vector Space. *Proceedings of the International Conference on Learning Representations (ICLR 2013)*, Scottsdale, AZ, 2-4 May 2013, 1-12.
- [17] Socher, R., Bauer, J., Manning, C.D. and Ng, A.Y. (2013) Parsing with Compositional Vector Grammars. *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, Sofia, Bulgaria, August 2013, 455-465.
- [18] Socher, R., Perelygin, A., Wu, J.Y., *et al.* (2013) Recursive Deep Models for Semantic Compositionality over a Sentiment Treebank. *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, Seattle, WA, October 2013, 1631-1642.
- [19] Rönqvist, S. and Sarlin, P. (2017) Bank Distress in the News: Describing Events through Deep Learning. *Neurocomputing*, **264**, 57-70.  
<https://doi.org/10.1016/j.neucom.2016.12.110>
- [20] Chang, W.L. and Wang, J.-Y. (2018) Mine Is Yours? Using Sentiment Analysis to Explore the Degree of Risk in the Sharing Economy. *Electronic Commerce Research and Applications*, **28**, 141-158. <https://doi.org/10.1016/j.elerap.2018.01.014>
- [21] Krishnamoorthy, S. (2018) Sentiment Analysis of Financial News Articles Using Performance Indicators. *Knowledge and Information Systems*, **56**, 373-394.  
<https://doi.org/10.1007/s10115-017-1134-1>
- [22] Tian, G.L., Si, Y., Qin, L. and Yu, Z.B. (2018) Internet Public Opinions, Response and Listed Firms' Information Efficiency. *Systems Engineering Theory & Practice*, **38**, 46-66.
- [23] Schumaker, R.P. and Chen, H. (2009) A Quantitative Stock Prediction System Based on Financial News. *Information Processing & Management*, **45**, 571-583.  
<https://doi.org/10.1016/j.ipm.2009.05.001>
- [24] Brown, G.W. and Cliff, M.T. (2004) Investor Sentiment and the Near-Term Stock Market. *Journal of Empirical Finance*, **11**, 1-27.  
<https://doi.org/10.1016/j.jempfin.2002.12.001>
- [25] Li, Q., Wang, T.J., Li, P., Liu, L., Gong, Q.X. and Chen, Y.Z. (2014) The Effect of News and Public Mood on Stock Movements. *Information Sciences*, **278**, 826-840.  
<https://doi.org/10.1016/j.ins.2014.03.096>
- [26] Weng, B., Lu, L., Wang, X., Megahed, F.M. and Martinez, W. (2018) Predicting Short-Term Stock Prices Using Ensemble Methods and Online Data Sources. *Expert Systems with Applications*, **112**, 258-273. <https://doi.org/10.1016/j.eswa.2018.06.016>
- [27] Hsu, F.M. and Liao, C.H. (2016) Does Information Uncertainty Moderate the Impact of Investors' Emotion on Stock Prices? 2016 *IEEE International Conference on Knowledge Engineering and Applications*, Singapore, 28-30 September 2016, 12-17.
- [28] Xu, Y.D. (2015) Knightian Uncertainty Emotion of Investors and the Huge Fluctuations of Stock Market. *Journal of Systems Engineering*, **30**, 736-745.
- [29] Shi, S.C., Zhu, Y.N., Zhao, Z.G., Kang, K.L. and Xiong, X. (2018) The Investor Sentiment Mined from WeChat Text and Stock Market Performance. *Systems Engineering Theory & Practice*, **38**, 1404-1412.
- [30] Lai, K.-S., Chen, H., Le, G.-A. and Dong, Y.-H. (2014) Can Mood Predict Stock Market? *Advances in Psychological Science*, **22**, 1770-1781.  
<https://doi.org/10.3724/SP.J.1042.2014.01770>
- [31] Gilbert, E. and Karahalios, K. (2010) Widespread Worry and the Stock Market. *4th*

---

*International AAAI Conference on Weblogs and Social Media*, Washington DC, 23-26 May 2010, 1-25.

- [32] Oh, C. and Sheng, O. (2011) Investigating Predictive Power of Stock Micro Blog Sentiment in Forecasting Future Stock Price Directional Movement. *Proceedings of the International Conference on Information Systems, ICIS 2011*, Shanghai, China, 4-7 December 2011, 57-58.
- [33] Tsibouris, G. and Zeidenberg, M. (1995) Testing the Efficient Markets Hypothesis with Gradient Descent Algorithms. In: *Neural Networks in the Capital Markets*, 127-136.
- [34] Wu, Q.-L. and Zhou, T.-H. (2016) Chinese Microblog Public Opinion Analysis Based on Topic Clustering and Emotional Intensity. *Information Studies: Theory & Application*, **39**, 109-112.
- [35] Zeng, R.X., Du, H.X. and Wang, J.Z. (2014) Comparative Study of Internet Public Opinion Index System, Method and Model. *Journal of Intelligence*, **33**, 96-101.
- [36] Zheng, X.X. and Chen, F.J. (2015) Research on the Evaluation Indicator System of Knowledge Complexity of the Network Public Opinion. *Modern Information*, **35**, 40-46.