Research on Key Technologies of Analysis of User Emotion Fluctuation Characteristics in Wireless Network Based on Social Information Processing

Jia Yu(✉)

School of Information Engineering, East China Jiaotong University, Nanchang 330013, China
liuwenjun0213@sina.cn

Abstract. With the rapid development of the mobile Internet, people's living habits have been greatly changed, and the ways to express their feelings have been gradually extended from offline to online. The research in the field of Internet user sentiment analysis provides a new method and tool for the prediction of social behavior in specific fields. Based on this, this research proposes the key technology of user's emotion fluctuation analysis in wireless network based on social information processing. The affective dictionary with multiple affective types is used to identify the affective types, and the database of users, products and affective types is formed. Based on large-scale general knowledge base, sentence level emotion is detected. Through the automatic learning of emotion expression, the classification of sentence emotion is realized, and the analysis model of network user comment emotion is constructed. The confusion matrix of the classifier is introduced to analyze the user's emotion fluctuation. In order to evaluate the effect of text emotion analysis, the accuracy index is introduced to evaluate the analysis results. Experimental results show that the proposed technique has ideal accuracy and strong applicability.

Keywords: Social information · Wireless network · Emotion fluctuation · Emotion characteristics

1 Introduction

In recent years, many scholars have used social media user sentiment analysis to study the prediction methods of social behavior in specific fields [1]. Since different social activities have different sensitivity to affective factors, affective analysis factors play different roles in different prediction tasks. A sentiment analysis method with independent predictive ability in a prediction task in a certain field may prove to not have independent predictive ability in another field, and can only be used as an important supplementary factor. In addition, the sentiment analysis methods used by different scholars are also different, which objectively affects the researcher's judgment on the importance of sentiment analysis.

S. Liu and X. Ma (Eds.): ADHIP 2021, LNICST 416, pp. 142–154, 2022.
https://doi.org/10.1007/978-3-030-94551-0_12

Literature [2] proposes a method for predicting user's emotional behavior intentions in message push process of mobile terminals, and constructs a correlation matrix to obtain the correlation between user's emotional behavior intentions and push messages. Based on the relational relationship, the evaluation function of user's emotional behavior intention is constructed, and the weights of user's emotional behavior safety are calculated by combining the evaluation function of user's emotional behavior intention. In reference [3], based on the theory of emotional load, with the help of coagmento system, 52 users were recruited to complete three planning tasks in a group of two. The relationship between task difficulty, time pressure, collaboration division, search results and collaboration results and users' negative emotion perception was investigated. However, the above two traditional methods ignore the statistical processing of emotional words, resulting in a large deviation in the analysis results of online users' emotional characteristics.

Based on this, this paper expounds the emotional analysis of network users, analyzes the key technologies of the emotional fluctuation characteristics of wireless network users based on social information processing, analyzes the typical methods and key technologies of applying emotional analysis to predict, and summarizes and prospects the existing problems and future development of current related research.

2 Key Technologies for Analyzing User Emotion Fluctuation Characteristics in Wireless Networks

2.1 Collection of Emotion Fluctuation Characteristics of Wireless Network Users

Sentiment classification refers to the automatic analysis of the relevant comments on goods, services, people and other research objects to find the reviewers' attitudes and opinions towards the research object. Online user sentiment analysis mainly focuses on automatic sentiment analysis of comment information generated by social media [4]. At present, the research in the field of online user sentiment analysis is mainly carried out in one of three levels: document level, sentence level and attribute level. According to the sentiment tendency of the whole review, the document level sentiment analysis divides the review into three polarity categories: positive, negative or neutral. It mainly uses various machine learning techniques to analyze the sentiment of reviews. A method used to classify opinions in multilingual web forums. Emotion analysis at sentence level mainly focuses on identifying subjective sentences and judging their emotional polarity [5]. Based on the method of sentiment dictionary, the sentiment lexicon with multiple sentiment type labels is used to identify sentiment types, and the database of users, products and sentiment types is formed. The method of emotion type discrimination is based on the intensity of emotion words in comparative sentences. Assuming that there are n emotion words in the corpus, the discriminant formula of emotion type is as follows

$$\text{SentiType } = \max\left(\sum_{i=0}^{n} D_i T_i\right) \tag{1}$$

Among them, SentiType is the emotional type; I is the sequence number of all emotion words involved in a comment, Di is the corresponding sequence number, the degree score of adverbs in the first five bytes of emotion words, and Ti is the emotion score of emotion words with corresponding sequence number [6]. In the case of considering degree adverbs, the emotional type corresponding to the emotional word with the highest emotional score in the review is selected as the overall emotional type of the sentence by ranking. Based on the technology of network projection, the calculation formula of the single-vertex network of the emotional network projection is proposed, such as the formula:

$$T = \text{SentiType} \times \left(\frac{Q^2}{N}\right) \tag{2}$$

Among them, Q: Is the network projection weight composed of commentator or comment object and emotion tag, and further proposes the method of emotion information extraction, and gives the experimental results and analysis of each method. The extracted emotion information includes three types: User dictionary, domain related emotion words and evaluation collocation. The extraction of evaluation collocation depends on domain related emotion words and user dictionary. For the construction of user dictionary, this paper proposes a method based on statistics, covering five kinds of statistics [7]. Secondly, co-occurrence graph is proposed to extract domain related emotional words. Finally, evaluation collocation extraction is realized based on word2vec model and syntax analysis. Experimental results show that the proposed method can accurately and efficiently extract the emotional information from the comment text. Based on this, this paper first analyzes the dependence of emotional information. The details are as follows (Fig. 1):

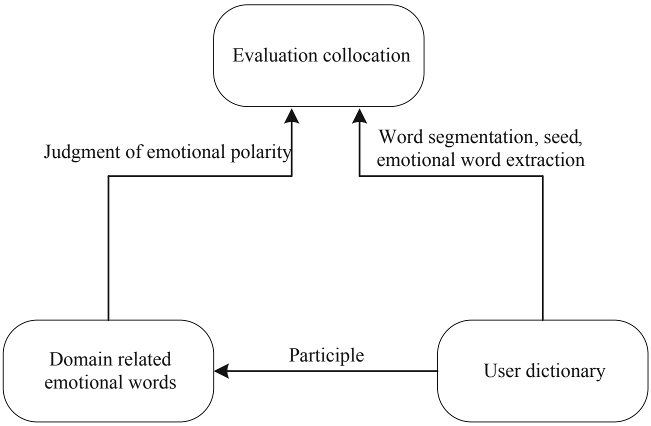

Fig. 1. Emotional information dependence

The emotional polarity of emotional words is not invariable, and the same emotional word may have polarity reversal when modifying different evaluation objects [8]. For example, the word “strong” is positive when modifying the atmosphere, and negative when modifying the smell of the room. The word “luxury” is negative in price decoration and positive in room decoration. In the task of sentiment analysis, it is very important to identify the emotion words whose polarity is not universal. This kind of emotion words is defined as "domain related emotion words". To solve this problem, a domain related emotion word recognition algorithm based on co-occurrence graph is proposed. "Co-occurrence graph" is a concept proposed, which refers to the undirected weighted graph based on the co-occurrence relationship of words in the corpus. The nodes in the graph are words in the corpus. If two words appear in the same sentence, they are connected.

2.2 The Relationship Between Interest and Passion of Wireless Network Users

Based on the large-scale general knowledge base to detect the sentence level emotion method, through the automatic learning of emotion expression to realize the classification of sentence emotion. The sentiment analysis of document layer and sentence layer is too rough to accurately determine the user's sentiment. To solve this problem, sentiment analysis at the attribute level, which extracts product-specific attribute opinions from reviews, is proposed. The part-of-speech tagging sequence rules are used to extract product attributes, and the polarity of the emotional description phrases of the attributes is determined based on the context information. In terms of emotional orientation, more related studies use positive and negative emotions to distinguish the emotional orientation in the text [9]. With the gradual deepening of research, some studies believe that such a simple emotional division may ignore many rich and multidimensional human emotional information. Therefore, some related studies try to further subdivide emotional orientation. In order to realize the organic combination of social information and emotion analysis, an online user comment emotion analysis model based on social information is constructed, as shown in Fig. 2.

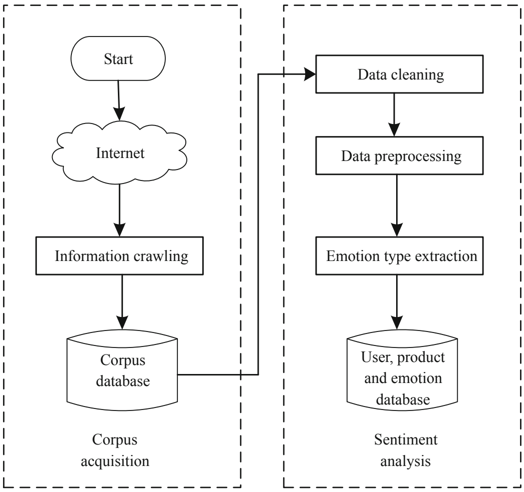

Fig. 2. Analysis model of online users' comments emotion

Sentiment analysis has been applied in many areas of prediction activities, and it has been proved that it can obtain better prediction results in most cases. However, due to the differences in application fields, emotion analysis and prediction target are not always closely related to each other [10]. In addition, sentiment analysis technology itself also has limitations in application fields or text processing objects, so in the research of prediction methods in different fields, sentiment analysis plays different roles. Let A be the behavior data set of social network users, and B be the data set to be recommended. The query history of a user is represented by a matrix.

$$S = T|A| \otimes |B| \tag{3}$$

Where, $a \in A$ represents the user, $b \in B$ represents the b in the set B, T represents the evaluation matrix, where each item $t_{ab} \in \{-1, 0, 1\}$ represents the search of user a for b, "–1" indicates that user a is not interested in b, "0" indicates that user a has not paid attention to b, "1" indicates that user a is interested in b.

Given a user's interest in $\{a, b\}$, let $Q\{a, b\}$ represent the degree of user a's interest in b, namely $Q \in \{A \otimes B \to T\}$. Given a data set list C to be recommended, it is expressed as follows:

$$R_a = \arg_{b \in C} topN\, SQ\{a, b\}, \forall a \in A \tag{4}$$

According to $Q\{a, b\}$, the score of each item in C is calculated, and then it is arranged in the order from big to small according to the score. The top N R_a with the highest score is recommended as the data cycle of the user's behavior. According to the different application ways of emotional factors in the prediction process, the prediction methods can be divided into the following two categories: The prediction method based on the results of emotional analysis as an auxiliary basis, and the prediction method based on the results of emotional analysis as the main basis. In the prediction method based on emotion analysis results, a combination of multiple indicators is used to predict. These indicators are usually proved to have a correlation with the predicted indicators, and the integration of multiple indicators can make the prediction effect optimal. Other prediction indicators integrated with the results of sentiment analysis include the number of mentions of the objects to be predicted in social media, the deformation of the number of mentions, and the historical data of authoritative institutions related to the research objects. Relevant studies believe that considering the emotional factors expressed by social media in the prediction process can effectively improve the prediction performance of existing indicators. The flow chart of typical prediction method based on the combination of emotion analysis results and other prediction indexes is shown in Fig. 3.

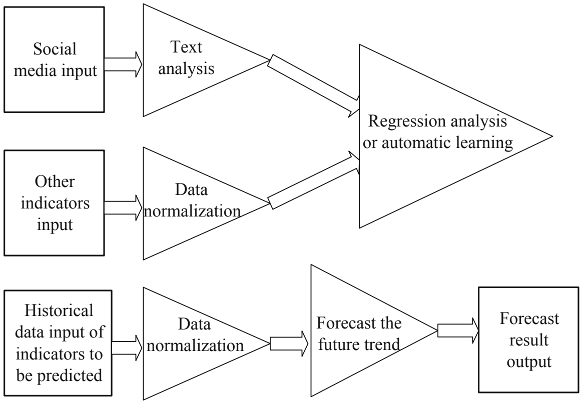

Fig. 3. Flow chart of emotion fluctuation analysis and prediction method

In the figure, the process of the prediction method based on the results of sentiment analysis mainly includes four key steps.

Text analysis of social media input. Text analysis can be divided into text preprocessing and text sentiment analysis. In the process of text preprocessing, it is necessary to standardize the social media content input in the form of natural text, so as to facilitate the subsequent text emotion analysis. The specific tasks of text preprocessing include: We should remove stop words, special characters other than exclamation marks and question marks, remove link addresses and user IDs, and use specific names to replace the names of the subjects (for example, the names of films with predicted box office revenue and books with predicted sales volume). In the process of text sentiment analysis, we can identify and count the sentiment words in the text with the help of the existing sentiment vocabulary to determine the text sentiment orientation. Machine learning method can also be applied to establish classification model through training to classify the emotion of the text. Because the third-party sentiment analysis tools are usually developed for general domain, it is difficult to make targeted adjustments when applied to specific domains, so the effect of sentiment analysis is not always satisfactory.

The input of other indicators and the input of historical data of indicators to be predicted are standardized. The input of other indicators to be integrated with the results of sentiment analysis and the historical data input of indicators to be predicted may have the problem of nonstandard data expression, which makes the subsequent indicators integration or comparison difficult. The range of data scale should be set, and the input of other indicators and historical data of indicators to be predicted should be standardized to a unified range of data scale.

Integrate the timing of sentiment indicators with the timing of other indicators. Regression analysis or automatic learning methods are used to integrate the standardized emotional indicator time series and other indicator time series to obtain the integrated indicator time series. Compared with the individual sentiment indicator timing and other indicator timings, the integrated indicator timing is more consistent with the indicator timing to be predicted when the timing advance is subtracted.

Use the integrated index trend to predict the future trend of the predictive index. The current integrated index trend reflects the future trend change of the index to be predicted, and the time advance n between the two can be calculated through the training text collection and historical index data to be predicted. Therefore, the trend of the integrated index near day t can be used to predict the trend of the index to be predicted on day t + n.

2.3 Realization of User Emotion Fluctuation Characteristic Analysis

Sentiment analysis has been applied in many areas of prediction activities, and it has been proved that it can obtain better prediction results in most cases. However, due to the differences in application fields, sentiment analysis and prediction target are not always closely related to each other. In addition, sentiment analysis technology itself also has limitations in application fields or text processing objects, so in the research of prediction methods in different fields, sentiment analysis plays different roles. According to the different ways in which affective factors are applied in the prediction process, the prediction methods can be divided into the following two categories: the prediction methods based on the results of affective analysis as the auxiliary basis, and the prediction methods based on the results of affective analysis as the main basis. The confusion matrix of classifier is introduced to explain the symbol, as shown in the Table 1:

Table 1. Confusion matrix and its symbol

Data category	Prediction example	Counter example of prediction
Practical examples	TP	FN
Actual counterexample	FP	TN

Different social media information sources have different update frequencies, and the performance of prediction methods based on these information sources is obviously affected by their update frequencies. Social network user behavior data cycle recommendation algorithm is carried out under the following premise, that is, if users score some items more similar, then they score other items more similar. Therefore, the first step of collaborative filtering is to establish a scoring matrix according to the user's interest in a certain item, that is, to score the user's interest items by using the user's browsing records, clicks, favorite content, etc. as the scoring index, and to establish a scoring matrix, as shown in Table 2 below.

Table 2. User project scoring matrix

Project	Project B1	Project B2	……	Project Bb	……	Project Bn
User A1	T11	T12	……	T1b	……	T1n
User A2	T21	T22	……	T2b	……	T2n
……	……	……	……	……	……	……
User Am	Tm1	Tm2	……	Tmb	……	Tmn

In the table, TMB is the score of user am on item BB, and a value in the table represents the score of a user on an item. Because wireless network has a limit of 140 characters, it is usually easier and more frequent to publish wireless network articles than blog articles. The timeliness of wireless network reflecting the emotional changes of network users is better than blog. Although the time advance of using wireless networks to predict epidemics is not obvious, shorter monitoring intervals, timely epidemic status information, and lower forecasting costs are fundamental to its existence.

The number of statements of related research objects in the information source should have a certain scale, and can more comprehensively and objectively reflect the category, proportion and intensity of people's emotional orientation towards research objects in the real world. Otherwise, the input of sparse data or false data will inevitably lead to the failure of prediction. The principle of richness of information source content and the principle of timeliness of information source update may need to make a trade-off under specific circumstances. For example, in the task of using sentiment analysis to predict the possible defects of automobile products, although wireless network information has the advantage of timely updating, its content length limit makes network users unable to describe the product experience and express emotions more deeply, so it is more suitable as a prediction information source.

Based on the quantitative rules of users' implicit emotion space, the corresponding relationships between the controllability, stability, transitivity, external force dependence of users on social platforms and the above characteristics are shown in the Table 3:

Table 3. Corresponding relationship between user characteristics and implicit emotion characteristics

User characteristics	Corresponding features
Controllability	Agree/disagree
Stability	Responsible/irresponsible, emotional/neurotic
Transitivity	Extroversion/introversion
External force dependence	Experience open/experience closed

As can be seen from the table, user controllability is measured by the "agree/disagree" feature. The more users tend to agree, the higher the controllability. For example, in a hot news event, if the user holds a negative view on the event, it can be used as a key monitoring and analysis object in public opinion control, and it needs to be properly guided and controlled by public opinion. User stability refers to the variability of user's emotional value for historical events, etc., which is measured by "responsible/irresponsible, emotional/neural sensitivity". User transitivity is measured by the "extroversion/introversion" feature. Outgoing users tend to have more attention, so they play an important role in the emotional communication of social networks. External force dependence refers to whether the user's emotion needs external force intervention. Experience open users need more external intervention to change their emotional tendency. This paper briefly introduces and discusses the method and model of text sentiment tendency analysis based on dictionary. Based on the emotion tendency calculation model of emotion group counting, the algorithm is designed and the emotion tendency of Chinese wireless network is analyzed. The main process of sentiment tendency analysis based on dictionary is shown in Fig. 4

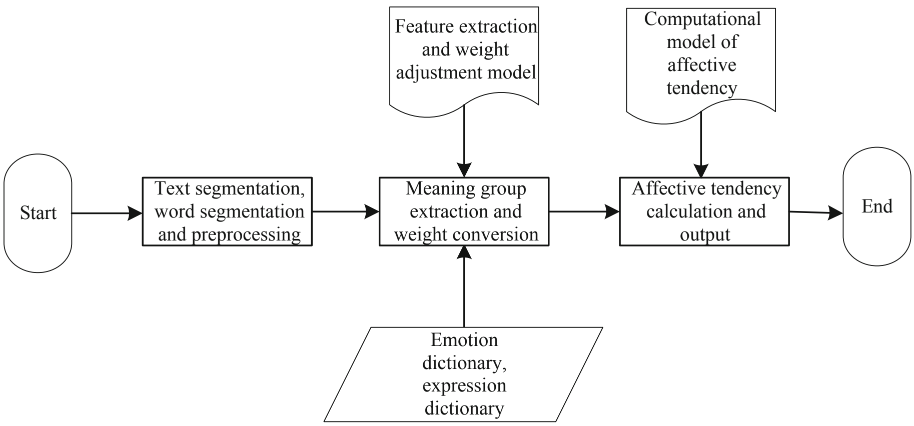

Fig. 4. Optimization of emotional tendency analysis process

Based on the above process, the analysis of user sentiment fluctuation can better guarantee the analysis effect, so as to effectively monitor the network environment.

3 Analysis of Experimental Results

In order to evaluate the effectiveness of the model, experiments and results analysis will be carried out according to the user emotion space model. Microblog is used as a test platform to query 1000 users' browsing records, clicks, and collections. Their interests and hobbies are classified as movie, food, beauty and movie stars. User project evaluation and scoring are conducted, and about 10 million scoring data are generated. As experimental data, social network user behavior data cycle is recommended (Table 4).

Use web crawlers to crawl wireless network user data. In order to ensure the timeliness of user data, only user data within half a year is extracted. According to the user

Table 4. Experimental environment configuration

Name	Parameter
CPU	Intel Core(TM)i5 -4590@3.30 GHz
Memory	4.00 GB
Operating system	64 bit Windows 8 professional
Development language	Python
Programming environment	Eclipse4.2
Compiler	JetBrains PyCharm Community Edition

emotional space construction steps in Sect. 5, user emotional characteristics are gradually established, three representative users are selected for detailed analysis, and how to analyze the emotional tendency of network users based on emotional space and user emotional characteristics. Here, X, D, and J are used as user nickname codes. Count the emotional frequency of three users. Emotional frequency refers to the total number of wireless networks sent by users in a single day, reflecting the frequency with which users express emotions. Figure 5 is a statistical comparison chart of the frequency trend of three users' emotions.

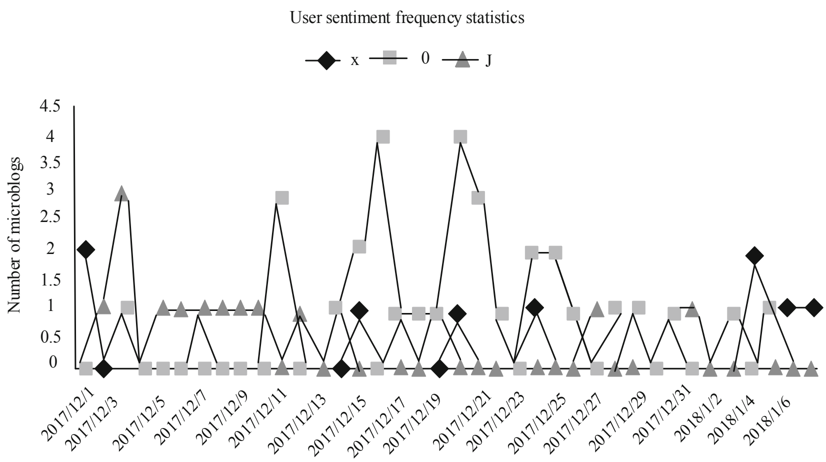

Fig. 5. Analysis of experimental results

Emotion density is the average daily emotion frequency value of users in a period of time, and the corresponding statistics of user emotion density are shown in the Table 5.

It can be seen from the table that the emotion density of user J is significantly higher than that of users X and D in the above time period, which indicates that user J is more inclined to express emotion in this time period. The experimental results of each eigenvalue based on implicit emotion space are discussed. It can be seen from the above table that the EI value of user x is much larger than the values of user D and J.

Table 5. Statistical table of emotional density of users

	X	J	D
Emotional density ω	0.447368421	0.868421053	0.421052632

according to the corresponding relationship between user characteristics and implicit emotion characteristics, it shows that the user has very strong emotion transitivity. By comparing their personal data, it is found that the number of fans of the user is 97,368,813, and the number of concerned users is 681. It really has a strong public opinion appeal and influence on the wireless network platform, and should be the key monitoring object of public opinion transmission guidance and supervision. Its D value is relatively large, which indicates that it has strong controllability, and the probability of needing external force interference is small, which indicates that the user's D emotion fluctuates greatly, the stability of emotion is low, and it may need to use external force to interfere with its emotional state when necessary. It shows that its emotional transitivity is relatively small, and the number of fans is only 136 through personal data check, which has very little influence on the wireless network platform and conforms to the quantitative judgment rules of transitivity. User J has a relatively large EI value. By comparing personal data, the number of fans is 1,492,514, and the number of concerned users is 1,277. It also has strong transitivity on the wireless network platform.

Based on the above experimental results, we test the accuracy of different methods. Methods of literature [2] and literature [3] were selected as experimental control group. The comparison results are shown in Fig. 6.

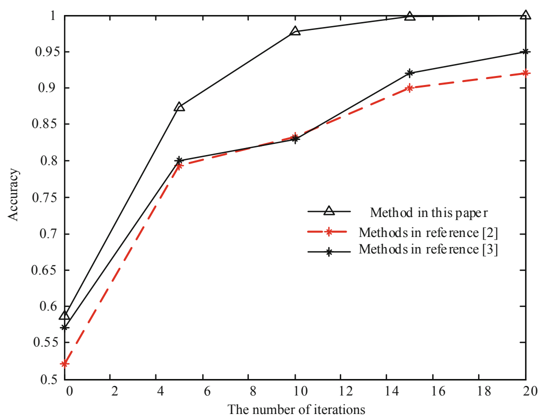

Fig. 6. Comparison of data mining accuracy

Analysis of Fig. 6 shows that with the increasing number of iterations, the accuracy of the user sentiment fluctuation analysis method proposed in this paper continues to improve, and the growth rate is fast. When the number of iterations is 4–12, the mining accuracy of literature [2] and literature [3] grows slowly, and the convergence is poor and the accuracy is low. Compared with the traditional method, the analysis of this method is more accurate and more applicable.

4 Conclusion and Prospect

Emotional expression is an innate instinct of human beings, which is the basis of human survival, mutual communication and social life. In recent years, the prediction based on Internet user sentiment analysis is an important field of sentiment analysis application research. At present, the typical prediction methods based on sentiment analysis can be summarized as the prediction methods based on sentiment analysis results and the prediction methods based on sentiment analysis results. In the process of prediction, the choice of emotion analysis source, the determination of prediction time advance and the statistical processing method of emotion words determine the real effect of prediction.

However, different social networks have different characteristics, for example, in anonymous social networks, anonymity will have an impact on emotional expression, the emotional expression of social networks used at different ages will have its own unique characteristics, these characteristics on the emotional expression of the impact can also be used as the content of the future study. In addition, the representation of user's emotion in different fields, the timely, accurate and comprehensive acquisition of corpus, and the correct analysis and statistics of user's emotion are the contents to be studied in the future.

Fund Projects. This work was supported by Jiangxi Education Department GJJ190316.

References

1. Faruolo, G., Santopietro, L., Saganeiti, L., Pilogallo, A., Scorza, F., Murgante, B.: The design of an urban Atlas to spread information concerning the growth of anthropic settlements in Basilicata Region. In: Gervasi, O., et al. (eds.) ICCSA 2020. LNCS, vol. 12255, pp. 214–225. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-58820-5_17
2. Wang, J.L., Luo, W.L.: Prediction of user information behavior intention in mobile terminal message push process. Comput. Simul. **36**(03), 440–443 (2019)
3. Huang, K., Yuan, X., Li, L., et al.: Research on users' negative emotion and related influence factors during collaborative information searching based on affective load theory. Doc. Inf. Knowl. **01**, 42–52 (2020)
4. Liu, S., Lu, M., Li, H., et al.: Prediction of gene expression patterns with generalized linear regression model. Front. Genet. **10**, 120 (2019)
5. Xu, X., Lin, J., Xiao, Y., et al.: An approach to generating the sequence of part variant design based on information transfer utility. Assembly Autom. **39**(1), 186–199 (2019)
6. Alegre Sepúlveda, T., Norambuena, B.K.: Twitter sentiment analysis for the estimation of voting intention in the 2017 Chilean elections. Intell. Data Anal. **24**(5), 1141–1160 (2020)
7. Bibi, M., Aziz, W., Almaraashi, M., et al.: A Cooperative binary-clustering framework based on majority voting for Twitter sentiment analysis. IEEE Access **8**(11), 68580–68592 (2020)
8. Fu, W., Liu, S., Srivastava, G.: Optimization of big data scheduling in social networks. Entropy **21**(9), 902 (2019)

9. Liu, S., Bai, W., Zeng, N., et al.: A fast fractal based compression for MRI images. IEEE Access **7**, 62412–62420 (2019)
10. Wu, T., Weld, D.S., Heer, J.: Local decision pitfalls in interactive machine learning: an investigation into feature selection in sentiment analysis. ACM Trans. Comput.-Hum. Interact. **26**(4), 1–27 (2019)