Research Article
Sentiment Analysis of English Text with Multilevel Features

Li Rao [1,2]

[1]*School of Foreign Studies, Nanjing University, Nanjing 210023, Jiangsu, China*
[2]*Jincheng College of Nanjing University of Aeronautics & Astronautics, Nanjing 211156, Jiangsu, China*

Correspondence should be addressed to Li Rao; raoli@nuaa.edu.cn

In natural language processing, text sentiment analysis is one of the important branches. It refers to the use of text mining and other technologies to extract attitudes, opinions, and other information from texts containing emotional information for analysis. Traditional sentiment analysis methods can be roughly divided into two categories: one is dictionary-based methods, and the other is machine learning-based methods. The former relies on the quality of the sentiment dictionary, while the latter relies on a large amount of high-quality data, so both have certain limitations. In text sentiment analysis research, word-level and sentence-level sentiment information extraction is a basic research task and has important research value. Through research, it is found that domain knowledge and context are two important factors influencing the extraction of emotional information. To this end, this paper proposes a text sentiment analysis method that integrates multiple features and constructs three features, which are based on the sentiment value feature of the dictionary, the expression feature, and the improved semantic feature, which are combined to build a text sentiment classification model. Aiming at the colloquial, irregular, and diverse features of English social media texts, this paper proposes a multilevel feature representation method. The sentiment classification experiments on English text show that the multilevel features proposed in this paper can effectively improve the F1_macro and accuracy of multiple model classifications. Compared with the existing research, the model in this paper improves the effect the most obvious.

1. Introduction

English is the world's largest language in Europe, America, Oceania, Asia, and other dozens of countries, and the total number of people who use it as a mother tongue or a second language is about hundreds of millions. English short texts with subjective emotional colors, summarizing, analyzing, and reasoning about the emotional information contained in them, are conducive to business decision-making, political public opinion analysis, and social trend prediction in relevant countries, and are useful for preventing precision political marketing, building harmonious and stable international relations, and advancing the "One Belt One Road" strategy of crossborder and inter-regional economic and trade, and cooperation and win-win cooperation are of great value. However, most of the current research in this field focuses on the strong position of English. There are not many researches in the field of English sentiment analysis, as shown in Figure 1. Some studies try to use English-related tools to obtain the results of sentiment analysis of English-English translations. However, due to the loss of sentiment and even semantics in the translation stage, the analysis stage ignores the characteristics of English itself, so the result is not ideal [1–10].

There are two major difficulties in the sentiment analysis of English social media texts: (1) the characteristics of English itself make free word order, polysemy, complex morphology, and nonprojection relationships that often exist in the text, which increases the complexity of semantic analysis and sentiment extraction; (2) social media texts have the characteristics of colloquialism, multiple slang words, irregular language, and unobvious context information.

The method based on sentiment dictionary mainly calculates based on the prior information of sentiment dictionary to judge the emotion contained in the text, but the size of sentiment dictionary is limited, and because of ignoring semantics, it is often impossible to get accurate classification. The machine learning method is based on the

Figure 1: English text sentiment analysis.

idea of pattern classification to deal with this problem. Through the artificial design of features, the text is vectorized and input into various classifiers for classification [11–15].

Dictionary-based sentiment analysis is one of the classic methods. It treats the text as a collection of words and ignores the connections between words. This method uses grammar and common expressions to set syntactic rules, builds an emotional dictionary manually, finally weights emotional words through matching emotional dictionaries and words, and counts the scores to determine the emotional polarity of the text. At present, there are many emotional dictionaries, most of which are hand-labeled. Commonly used Chinese emotional dictionaries include HowNet emotional dictionaries. Thelwall et al. proposed the classic algorithm SentiStrength based on the sentiment dictionary, which adjusted the method of calculating the sentiment value of social networks and achieved good results. Some scholars think that sentiment analysis cannot be discussed without leaving the context, so they proposed an algorithm to explore the connection between modifiers and sentiment words, which is more effective in predicting the sentiment polarity of the text. Some scholars believe that ordinary sentiment dictionaries are the shortcomings of sentiment analysis. They supplemented the original dictionaries, seized most of the short content features in Weibo texts, and built a new dictionary on sentiment analysis with the help of new weighting rules algorithms. Although many researchers continue to improve and supplement emotional dictionaries and have made some progress, this method is always limited by the dictionary, it is impossible to include all emotional words, and it cannot adapt to the times. In addition, because it is not suitable for texts with implicit sentiment characteristics, the accuracy of this method has not been high when used in text sentiment analysis [16–19].

Sentiment analysis methods based on machine learning are usually not limited by dictionaries and are mainly used for model training. The general steps are as follows: prepare manually labeled text data for training an emotional classification model, optimize parameters, and finally predict the emotional tendency of unknown data through the model. Pang et al. used Naive Bayesian and support vector machine (SVM) algorithms for the first time in the field of sentiment analysis in 2002 and studied the related comments of movies as the object of sentiment classification. Significant progress has been made. When analyzing English reviews, Beineke

et al. combined machine learning algorithms with manual text annotations and obtained excellent results. Then, Rao et al. proposed a model, the LDA topic generation model, which uses the bag-of-words method to recognize text or corpus. Mou et al. used three algorithms: Naive Bayes, K-nearest neighbor, and SVM in the model to judge the emotional tendency of English text. Dey et al. compared the two algorithms of Naive Bayes and K-nearest neighbors, and conducted experiments on hotel and movie review text sets [20–25]. The results showed that the effect of Naive Bayes is better. Mukras et al. proposed a method that can learn and mark part of speech from a text library. Mertiya et al. combined machine learning methods with dictionaries and performed sentiment analysis around comments on Facebook, providing a new idea of combining the two. Firstly, the features are extracted through the sentiment dictionary, and then, the text sentiment polarity is judged by Naive Bayes. In sentiment analysis, most of the methods based on machine learning are based on statistical theory, alleviating the problem of dictionaries that consume a lot of manpower and time and at the same time improving the accuracy of sentiment analysis, but there are still some shortcomings, such as the quality of data required for feature extraction—good and large numbers. In the current era of big data, its efficiency needs to be improved [26–29].

On the whole, machine learning methods perform better than rule methods. However, for complex English, traditional machine learning modeling methods cannot achieve satisfactory results. To this end, this paper proposes a text sentiment analysis method that integrates multiple features and constructs three features, which are based on the sentiment value feature of the dictionary, the expression feature, and the improved semantic feature, which are combined to build a text sentiment classification model.

2. Text Sentiment Analysis Combining Multiple Features

Due to the limited resources of English sentiment analysis and its unique complexity, it has become a challenging task to identify the sentiment of English comments. This paper proposes a multifeature fusion English text sentiment analysis method that combines machine learning and sentiment rules. The goal is to classify the sentiment of the existing review text, so as to find the user's evaluation information on products and topics. The emotional results are mapped to polarity. The flowchart of text sentiment analysis fused with multiple features is shown in Figure 2.

2.1. Building an Emotional Dictionary. The sentiment dictionary is a set of color markers with sentiment polarity. It is an indispensable part of the text sentiment analysis task. Generally, the more complete the sentiment dictionary, the more accurate the recognition results will be. In order to get better recognition results, the currently widely used emotional dictionaries (such as HowNet, NTUSD, and TSING) are integrated and expanded,
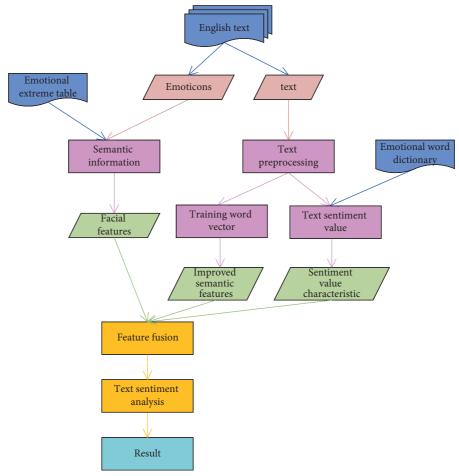
FIGURE 2: Flowchart of text sentiment analysis fused with multiple features.

and the basic emotional words, expression emotional words, degree adverbs, negative words, and transitional conjunctions have been established—comprehensive emotional dictionary. In addition, a dictionary of online emotional words has also been established. For the emergence of new words on the Internet, many documents have studied the method of expanding the emotional dictionary based on machine learning and have achieved certain results. However, due to the problems of word segmentation and candidate word selection, algorithms cannot be used to get a good processing effect for all kinds of network terms that are emerging in an endless stream. Therefore, based on the network term dictionary crawled by Zhihu, other network emotion words were sorted and expanded, and a network emotion dictionary with 726 emotion words was constructed. Dictionary-based sentiment value features refer to constructing specific rules based on sentiment dictionary and modifier dictionary, matching the sentiment words and modifiers contained in the text, and then performing weighting calculations to obtain sentiment value features as the expression form of text emotion. To calculate the sentiment value of the text, the formula is as follows:

$$\text{Score} = \sum_{j=1}^{m} \text{base} \times \prod_{i=1}^{n} \text{weight}_i. \tag{1}$$

Among them, m is the total number of emotional words contained in the text, n is the number of modifiers of a certain emotional word, base is the basic score, and weight is the degree adverb or the weight of the negative word.

2.2. Text Preprocessing. English text usually contains a strong personal style and personal emotional color, and the expression content is rich. In addition to being irregular, the grammar is basically biased toward daily life and colloquialism, but also contains a large number of irregular language, typos, links and expressions, symbols, etc., so it needs to be preprocessed before proceeding with the task of text sentiment analysis. In order to improve the efficiency of text sentiment analysis, the first step is to filter out URLs, tags, and irregular language and remove stop words. In the stage of text preprocessing, word segmentation is one of the very important components. Because the comment text has obvious colloquial characteristics and contains a large number of new words on the Internet, the effect of using general word segmentation tools is not very satisfactory. Therefore, the English word segmentation system ICTCLAS developed by the Chinese Academy of Sciences that can be added to user-defined dictionaries is used to segment the comment text to achieve better word segmentation effect.

2.3. Dictionary-Based Emotional Rule Classification Method.
The most classic is to accumulate the sentiment words to obtain the sentimental value of the text. The formula is as follows:

$$S = \sum_{i=1}^{n} Sw_i. \tag{2}$$

Among them, Swi is the polarity of the i-th emotional word and n is the total number of emotional words.

According to formula (2), the polarity of all sentiment words is superimposed, and the sentiment tendency value of the text is judged according to the finally obtained value. However, it is not only the emotional words that determine the emotional polarity in the text. Others such as negative words, degree adverbs, and language structure will all have a certain impact on the emotional tendency. Aiming at the shortcomings of the classic methods, a dictionary-based emotional rule classification method is proposed. Since the review text is generally short, first, each clause in the text is used as a unit, the emotion calculation formula (3) obtained by the emotion rule method based on the emotion dictionary is calculated for each unit, and finally all the units are calculated. The scores are superimposed to get the emotional orientation of the entire review text.

$$S_{\text{unit}} = \sum_{i=1}^{n} \left(K * Pw_i * \prod_{j=1}^{m} \text{mod}_j \right). \tag{3}$$

Among them, n represents the total number of emotional words in the text; Pw_i represents the extreme value of the i-th emotional word; m represents the number of words that modify the i-th emotional word; modj represents the weight of its corresponding modifier; and k represents the strengthening and weakening coefficient, to avoid sentiment analysis bias caused by subject confusion. In various algorithms of text sentiment analysis tasks, often due to the lack of referential judgment, the sentiment polarity obtained is not the judgment of the subject, and the results are biased. The overall architecture of the MFCNN model is shown in Figure 3.

Extract the emoticons in the text and calculate the extreme value of text emotion as follows:

$$\text{Score} = \frac{\sum_{i=1}^{m} F(e_i, \text{pos})}{m} + \frac{\sum_{j=1}^{m} F(e_j, \text{neg})}{n}, \tag{4}$$

where m and n are the number of positive emoticons and negative emoticons in the text, e is emoticons, pos and neg are the extreme value tables of positive and negative emoticons, and the function of F is to take out the scores of corresponding emoticons in the extreme value. The x with different y is shown in Figure 4.

It is convenient to intuitively understand the relationship between the number of expressions in the text and the extreme tendency of the text, and the cumulative distribution function (CDF) is introduced. The definition formula is as follows:

$$F_X(x) = P(X \le x). \tag{5}$$

2.4. Improved Semantic Features. The text word vector is regarded as the semantic feature of the text, because it contains the semantic information of the word, it is regarded as the semantic feature of the text. The Word2vec model is used to convert the text into word vectors, which alleviates the problems of sparse matrix and excessive dimension, and retains the sequence information of the words in the text, but omits the different importance of different words to the text. The TF-IDF algorithm just solves this problem, so the TF-IDF and Word2vec are combined, and the text word vector trained by the model is called the semantic feature of text improvement. It combines the advantages of the two, not only retains the sequence information of the words in the text, but also gives different weights to different words in the text. Assuming a text *di*, the number of words after word segmentation is M, and the word vector dimension is N; the text is expressed as follows:

$$d_i = \langle w_1, w_2, \cdots, w_M \rangle. \tag{6}$$

The word vector is generated through the Word2vec model. The text contains multiple words, and each word has its corresponding word vector. Join them to obtain the $M \times$ N-dimensional vector matrix of the text:

$$G(d_i) = \text{matrix}(M \times N). \tag{7}$$

Multiplying with the weight matrix is the vector matrix obtained by the improved Word2vec:

$$W\_G(d_i) = \text{matrix}(M \times N). \tag{8}$$

The expression formula is as follows:

$$G(d_i) = \{W2v(w_1), W2v(w_2), \cdots, W2v(w_M)\},$$
$$W\_G(d_i) = \{\text{weight}(w_1)W2v(w_1), \cdots, \text{weight}(w_M)W2v(w_M)\}. \tag{9}$$

Among them, each vector $W2v(wi)$ in the G(di) vector matrix is the word vector of the word wi in the text, which is obtained by training the Word2vec model; each vector in the $W - G(di)$ vector matrix weight $(wi)v(wi)$, where weight (wi) is the weight value of the word wi calculated by the TF-IDF algorithm; and multiplying weight (wi) and $W2v(wi)$ is the word vector of the improved Word2vec, which is the word vector of each word in the text. The composed text vector matrix $W - G(di)$ is used as the improved semantic feature of this article. The predicted value is compared in Figure 5.

2.5. Method Based on Machine Learning. The classification method based on machine learning regards sentiment analysis as a pattern classification problem and establishes a classification model to judge sentiment polarity. First, the machine learning method needs to label the text, use it as a training set, then extract features to train the classifier, and finally test the test corpus to get the classification result. Text
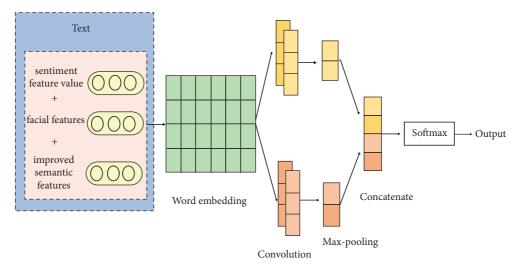
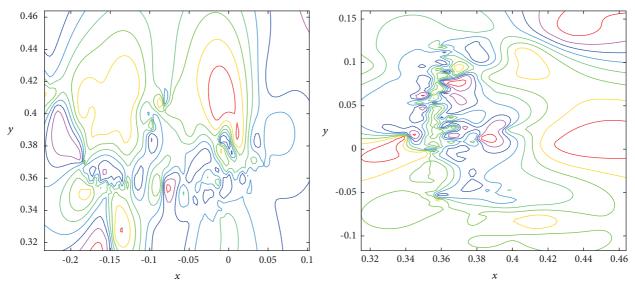FIGURE 3: The overall architecture of the MFCNN model.

FIGURE 4: y vs. x.

feature selection is a key step of machine learning, which determines the accuracy of sentiment classification. Three types of features are selected in the article: unigram features, syntactic features, and dependent word collocation features. Among them, the syntactic feature is the feature of the research component and the order of arrangement. Considering that the phrase structure can reduce sentence ambiguity, the binary word (bigram) and its combined part-of-speech tag are added to the feature set as its feature; the dependency relationship feature is from the dependency parse tree. The dependence relationship identification obtained plays an important role in the labeling of emotional category information, and it can save the information directly related to emotional words and other hidden information. The amplitude variation is shown in Figure 6.

From the above analysis steps, three basic feature templates of the machine learning method can be obtained. In order to avoid the problem of the degradation of the classifier effect due to the large dimensionality of the original feature space, the feature selection method of information gain (IG) is used to reduce the dimensionality of the original feature space to select the corresponding features. The formula is as follows:

$$\text{IG}(T) = H(C) - H(C|T),$$

$$H(C) = \sum_{i=1}^{n} P(C_i)\log_2(C_i). \quad (10)$$

Among them, $P(C_i)$ represents the probability of category C_i.

2.6. Text Sentiment Analysis Method Fused with Multiple Features.

The algorithm that combines machine learning methods and rule methods has attracted the attention of many researchers. For example, Qiu et al. used the dictionary
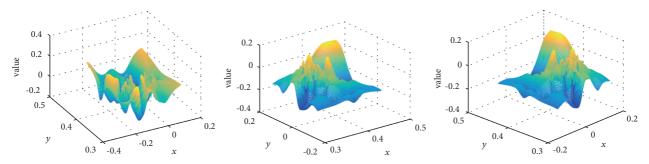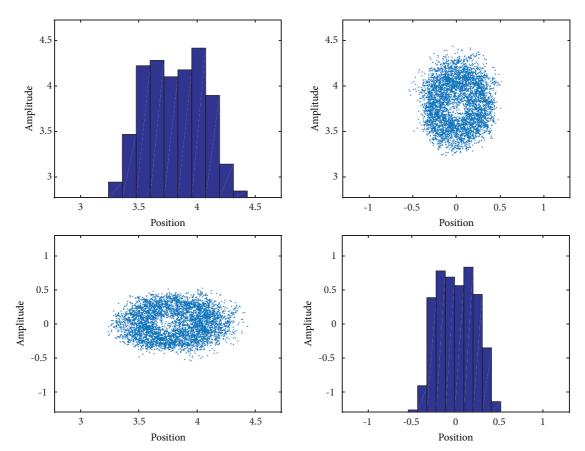
Figure 5: Predicted value comparison.

Figure 6: Amplitude in different positions.

classification results as the training corpus of the classification model to form a hierarchical iterative classification framework; Mohammad et al. accumulated emotional words and ended words—the polarity as a feature. Inspired by the predecessors, this paper proposes a multifeature fusion classification algorithm that combines machine learning and emotional rules. As a necessary step of the machine learning and emotional rule fusion method, after calculating the emotional score based on the improved emotional rule method, the effective information is extracted and expanded to integrate with machine learning features. Four characteristics are extracted in the article: emotional word score, the ratio of the number of positive/negative emotional words, the ratio of the number of enhancements to the number of weakenings, and the ratio of the number of praise/derogation emotional sentences, which are normalized and expanded to machine learning feature templates. Train the SVM classifier, and then use the test corpus to test. Through the above process, a multifeature fusion text classification method combining machine learning method and dictionary-based emotional rule method is realized, and multiple effective emotional information extracted from the rule algorithm is extended to the vector space, so that the machine learning algorithm can make full use of the characteristics of the rules and learn more emotional knowledge. The prediction is compared in Figure 7.

Emotional features play an important role in sentiment classification tasks. Its extraction is closely related to the
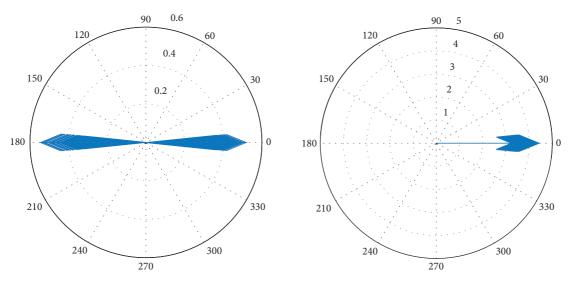
Figure 7: Prediction.

quality of the classification results. This paper constructs three kinds of emotional features: based on dictionary-based emotional value features, expression features, and improved semantic features. The feature-based text sentiment analysis method is to splice multiple features to form a multifeature vector matrix, which is used as the input of the convolutional neural network. This method can extract multidimensional sentiment information.

First, divide the text, store the text part in Dt, and store the expression part in De. Perform text preprocessing on Dt, and combine the sentiment dictionary and modifier dictionary to calculate the sentiment value characteristics of the text. Dt is trained through the improved Word2vec model to obtain the text word vector, which constitutes an improved semantic feature. De combines the emotional extreme value table of emoticons to calculate the emotional extreme value of the expression, plus the number of appearances of the expression and semantic information, which together constitute the expression feature. The three characteristics are merged to perform text sentiment analysis. TextCNN is one of the popular deep learning models. It is adjusted on the basis of CNN to make the TextCNN model more suitable for extracting text features and is often used in sentiment analysis. This chapter uses it as the core model and proposes an emotional classification model MFCNN based on multifeature fusion, which converts different features into corresponding vectors, uses splicing for feature fusion, constructs a multifeature vector matrix, and inputs it into the text convolutional neural network. Finally, get the classification result.

3. Experiment and Analysis

Research on sentiment analysis often requires a lot of text as support to facilitate training models. However, the amount of relevant text data on the Internet is not many and the quality is low, so this article crawled 10,000 texts through

crawlers and labeled them as positive emotions or negatives according to their sentiment tendencies. Emotions constitute a two-category data set. Among them, there are 5497 positive sentiment texts and 4503 negative sentiment texts, and the sentiment tendencies are roughly balanced between positive and negative. The overall process of the crawler is roughly as follows: randomly take out an account from the account pool of the database to simulate the login of a Weibo user, obtain cookie information, so that the website can identify the user's identity, obtain the target URL resource, use the Python Requests Library to process HTTP requests, return a response object, then use the BeautifulSoup4 library to parse and process HTML, and finally use regular expressions to extract text data and save it to the database. The evaluated data are shown in Figure 8. As can be seen, the data and value vary in each condition, which shows the validation of the proposed method.

The text preprocessing is divided into three steps: data cleaning, Chinese word segmentation, and stop word removal, which are described in detail below. The first step is data cleaning. Data cleaning is to remove characters and data that have nothing to do with the text content, such as URL links, forwarding symbols //, topic symbols #, designated user symbols @, and other information in the text. This information has nothing to do with the content of the text, but it may interfere with the results of sentiment analysis and affect subsequent word segmentation. For these data, this article uses regular expressions to delete them all. The method of replacing Weibo emoticons is adopted. In order to facilitate the construction of emoticon features, the emoticons contained in the text are converted into "[emoticons]" format when crawling the Weibo text. For example, the emoticon that represents disappointment is transformed into "[disappointment]" in the text, and the emoticon is placed in "[ ]" to distinguish it from the text.

Text vectorization is to convert each word into a vector after preprocessing and dividing the word, and then, each
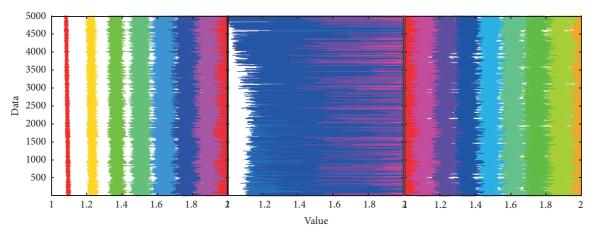
FIGURE 8: Evaluated data.

word vector is formed into a vector matrix according to the order of the words in the text. In this way, mapping the words to the vector space preserves its semantic information. Text vectorization is the cornerstone of text research. Whether or not the word vector can be expressed correctly will affect the judgment of text orientation. This paper uses the improved Word2vec model to vectorize the text and compares the final sentiment classification results. The specific configuration of the experiment is as follows: the processor is Intel(F)Core(TM)i5-6500CPU@3.2 GHz; the memory is 8 GB; the programming platform is Eclipse; the development language is Java; the database is SqlSever2008.

In order to verify the effectiveness of the text sentiment analysis method of fusion of multiple features, this section designs 7 models to conduct comparative experiments on the data set. The 7 models use different feature construction methods to form a vector matrix as the input of TextCNN and compare the final text—the result of sentiment classification. These seven models are as follows: (1) CNN model: Word2vec model trains word vectors, that is, the original semantic features, and then inputs them to TextCNN for text sentiment classification. (2) TCNN model: The word vector is trained by the Word2vec model weighted based on the TF-IDF algorithm, that is, the improved semantic features, and then input to TextCNN for text sentiment classification. (3) SCNN model: On the basis of the CNN model, fusion of emotional value features is based on dictionary, that is, the fusion of original semantic features and emotional value features. (4) ECNN model: On the basis of the CNN model, expression features are fused; that is, the original semantic features and expression features are fused. (5) TSCNN model: On the basis of the TCNN model, fusion of emotional value features is based on dictionary, that is, the fusion of improved semantic features and emotional value features. (6) TECNN model: On the basis of the TCNN model, the expression features are integrated; that is, the improved semantic feature and the expression feature are fused. (7) MFCNN model: Fusion of multiple features, fusion of dictionary-based emotional value features, expression features, and improved semantic features form a multifeature vector matrix, which can extract multidimensional

emotional information as the input of TextCNN for emotional classification.

The following conclusions can be drawn: (1) Compared with the CNN model with only original semantic features, the accuracy of the TCNN model with improved semantic features is increased by 2.19%, indicating that the combination of TF-IDF algorithm improves the keyword in the text. The weight is beneficial to the improvement of sentiment classification performance and verifies the effectiveness of the improved semantic features. (2) The ECNN model, which integrates expression features on the basis of original semantic features, has an accuracy rate of 1.98% higher than that of the CNN model, which shows that emojis increase the effect of indicating emotions, and also proves the necessity of adding expression features. In summary, it can be seen that the improved semantic features and expression features do contain more emotional information that is conducive to classification for the self-built data set. (3) Compared with the CNN model, the accuracy of the SCNN model fused with emotional value features based on the original semantic features is increased by 0.53%, and the improvement effect is not very significant. It can be seen that the fusion of a single emotional value feature has limited improvement for the model. (4) The TSCNN model that combines improved semantic features and emotional value features has an accuracy rate of 2.42% higher than that of the CNN model. The TECNN model that combines improved semantic features and expression features has an accuracy rate of 3.17% higher than that of the CNN model. The accuracy of classification is improved. It has been further improved; the accuracy rate of the MFCNN model that finally fused the three features is 0.59% to 3.76% higher than other comparison models, indicating that the MFCNN model can learn more dimensional emotional information of the text from the vector matrix fused with multiple features. Prove the feasibility and effectiveness of this method.

Compared with the CNN model, the MFCNN model fused with multiple features has nearly 5% and 4% improvement in accuracy and F1 value, respectively, and an improvement in recall rate of 2%. Among them, the accuracy of the TCNN model and the MFCNN model has been significantly improved, and it can be seen that the improved semantic features

perform better. Regarding the recall rate, the ECNN model has achieved the best results, which is 4% higher than the CNN model. On the F1 value, comparing the results of the CNN model and the SCNN model, and the TECNN model and the MFCNN model, it can be seen that the individual emotion value feature has a small effect on the improvement of the F1 value, but the multifeatured model has achieved excellent results. This paper proposes a Word2vec model training word vector based on the TF-IDF algorithm. In order to prove its effectiveness, it is compared with the traditional Word2vec model in three sets of experiments.

4. Conclusion

In natural language processing, text sentiment analysis is one of the important branches. It refers to the use of text mining and other technologies to extract attitudes, opinions, and other information from texts containing emotional information for analysis. Traditional sentiment analysis methods can be roughly divided into two categories: one is dictionary-based methods, and the other is machine learning-based methods. The former relies on the quality of the sentiment dictionary, while the latter relies on a large amount of high-quality data, so both have certain limitations. In text sentiment analysis research, word-level and sentence-level sentiment information extraction is a basic research task and has important research value. Through research, it is found that domain knowledge and context are two important factors influencing the extraction of emotional information. To this end, this paper proposes a text sentiment analysis method that integrates multiple features and constructs three features, which are based on the sentiment value feature of the dictionary, the expression feature, and the improved semantic feature, which are combined to build a text sentiment classification model. Aiming at the colloquial, irregular, and diverse features of English social media texts, this paper proposes a multilevel feature representation method. The sentiment classification experiments on English text show that the multilevel features proposed in this paper can effectively improve the F1_macro and accuracy of multiple model classifications. Compared with the existing research, the model in this paper improves the effect the most obvious.

Data Availability

The data set can be accessed upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

[1] Y. Yu, C. Yang, Q. Deng, T. Nyima, S. Liang, and C. Zhou, "Memristive network-based genetic algorithm and its application to image edge detection," *Journal of Systems Engineering and Electronics*, vol. 32, no. 5, pp. 1–9, 2021.

[2] Y. Ishida and S. Hashimoto, "Asymmetric characterization of diversity in symmetric stable marriage problems: an example of agent evacuation," *Procedia Computer Science*, vol. 60, no. 1, pp. 1472–1481, 2015.

[3] P. Zoha and R. Kaushik, "Image edge detection based on swarm intelligence using memristive networks," *IEEE Trans. on CAD of Integrated Circuits and Systems*, vol. 37, no. 9, pp. 1774–1787, 2018.

[4] J. Pais, "Random matching in the college admissions problem," *Economic Theory*, vol. 35, no. 1, pp. 99–116, 2018.

[5] J. J. Jung and G. S. Jo, "Brokerage between buyer and seller agents using constraint satisfaction problem models," *Decision Support Systems*, vol. 28, no. 4, pp. 291–384, 2020.

[6] Y. Liu and K. W. Li, "A two-sided matching decision method for supply and demand of technological knowledge," *Journal of Knowledge Management*, vol. 21, no. 3, JKM-05-2016-0183, 2017.

[7] J. Byun and S. Jang, "Effective destination advertising: matching effect between advertising language and destination type," *Tourism Management*, vol. 50, no. 10, pp. 31–40, 2015.

[8] A. N. Nagamani, S. N. Anuktha, N. Nanditha, and V. K. Agrawal, "A genetic algorithm-based heuristic method for test set generation in reversible circuits," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 37, no. 2, pp. 324–336, 2018.

[9] C. Koch and S. P. Penczynski, "The winner's curse: c," *Journal of Economic Theory*, vol. 174, pp. 57–102, 2018.

[10] C. K. Karl, "Investigating the winner's curse based on decision making in an auction environment," *Simulation & Gaming*, vol. 47, no. 3, pp. 324–345, 2016.

[11] D. Ettinger and F. Michelucci, "Creating a winner's curse via jump bids," *Review of Economic Design*, vol. 20, no. 3, pp. 173–186, 2016.

[12] J. A. Brander and E. J. Egan, "The winner's curse in acquisitions of privately-held firms," *The Quarterly Review of Economics and Finance*, vol. 65, pp. 249–262, 2017.

[13] Z. Palmowski, "A note on var for the winner's curse," *Economics/Ekonomia*, vol. 15, no. 3, pp. 124–134, 2017.

[14] B. R. Routledge and S. E. Zin, "Model uncertainty and liquidity," *Review of Economic Dynamics*, vol. 12, no. 4, pp. 543–566, 2009.

[15] D. Easley and M. O'Hara, "Ambiguity and nonparticipation: the role of regulation," *Review of Financial Studies*, vol. 22, no. 5, pp. 1817–1843, 2019.

[16] P. Klibano, M. Marinacci, and S. Mukerji, "A smooth model of decision making under ambiguity," *Econometrica*, vol. 73, no. 6, pp. 1849–1892, 2005.

[17] Y. Halevy, "Ellsberg revisited: an experimental study," *Econometrica*, vol. 75, no. 2, pp. 503–536, 2017.

[18] D. Ahn, S. Choi, D. Gale, and S. Kariv, "Estimating ambiguity aversion in a portfolio choice experiment," *Working paper*, vol. 5, no. 2, pp. 195–223, 2019.

[19] T. Hayashi and R. Wada, "Choice with imprecise information: an experimental approach," *Theory and Decision*, vol. 69, no. 3, pp. 355–373, 2010.

[20] K. Zima, E. Plebankiewicz, and D. Wieczorek, "A SWOT analysis of the use of BIM technology in the polish construction industry," *Buildings*, vol. 10, no. 1, p. 16, 2020.

[21] P. Sun, B. Liu, and T. Sun, "Injury status and strategies of female 7-a-side rugby players in Anhui Province," *Sports Boutique*, vol. 38, no. 03, pp. 72–74, 2019.

[22] P. Guild, M. R. Lininger, and M. Warren, "The association between the single leg hop test and lower-extremity injuries in female athletes: a critically appraised topic," *Journal of Sport Rehabilitation*, vol. 30, no. 2, pp. 1–7, 2020.

[23] U. G. Inyang, E. E. Akpan, and O. C. Akinyokun, "A hybrid machine learning approach for flood risk assessment and classification," *International Journal of Computational Intelligence and Applications*, vol. 19, no. 2, Article ID 2050012, 2020.

[24] Q. Liu, S. Du, B. Wyk, and Y. Sun, "Double-layer-clustering differential evolution multimodal optimization by speciation and self-adaptive strategies," *Information Sciences*, vol. 545, no. 1, pp. 465–486, 2021.

[25] H. R. Medeiros, F. D. Oliveira, H. F. Bassani, and A. Araujo, "Dynamic topology and relevance learning SOM-based algorithm for image clustering tasks," *Computer Vision and Image Understanding*, vol. 179, no. FEB, pp. 19–30, 2019.

[26] Y. Deng, D. Huang, S. Du, G. Li, and J. Lv, "A double-layer attention based adversarial network for partial transfer learning in machinery fault diagnosis," *Computers in Industry*, vol. 127, Article ID 103399, 2021.

[27] J. J. Chan, K. K. Chen, S. Sarker et al., "Epidemiology of Achilles tendon injuries in collegiate level athletes in the United States," *International Orthopaedics*, vol. 44, no. 3, pp. 585–594, 2020.

[28] W. Li, G. G. Wang, and A. H. Gandomi, "A survey of learning-based intelligent optimization algorithms," *Archives of Computational Methods in Engineering*, vol. 28, pp. 3781–3799, 2021.

[29] G. G. Wang, A. H. Gandomi, A. H. Alavi, and D. Gong, "A comprehensive review of krill herd algorithm: variants, hybrids and applications," *Artificial Intelligence Review*, vol. 51, no. 1, pp. 119–148, 2019.