

Sentiment Analysis Technology of Chinese Weibo under the Background of Big Data

Ruwei Wang

Shandong University of Finance and Economics, Jinan 250000, China

Abstract. Chinese Weibo sentiment analysis technology is an important driving force for the development and progress of natural language processing, machine learning and other related fields. The purpose of this paper is to discuss the research of Chinese Weibo's sentiment analysis technology under the background of big data. At first, it summarizes the development history and current situation of sentiment analysis technology, then introduces the development status and challenges of Chinese Weibo's sentiment analysis technology, and finally looks forward to the future development of Chinese Weibo's sentiment analysis technology. At the same time, this paper deeply discusses the challenges faced by Chinese Weibo sentiment analysis technology, such as the processing of massive data, the processing of unstructured data, the complexity of sentiment understanding and the diversity of languages. In view of these challenges, this paper puts forward corresponding solutions and suggestions, and provides a useful reference for the development of Chinese Weibo sentiment analysis technology. Generally speaking, the emotional analysis of Chinese Weibo in this paper has important practical significance and application value, with a view to providing valuable theoretical support and practical guidance for the research and practice in related fields.

Keywords: Big Data; Deep Learning; Chinese; Weibo's Emotional Analysis.

1. Introduction

In the context of big data, sentiment analysis technology, as an important natural language processing method, has attracted more and more attention and application [1]. Sentiment analysis aims to automatically identify and understand the emotions and opinions expressed in the text, and has important guiding significance for many fields such as market research, public opinion monitoring, decision-making and so on [2]. However, the Chinese Weibo sentiment analysis technology faces many challenges in practice, such as the processing of massive data, unstructured data, the complexity of emotional understanding and the diversity of languages [3]. With the popularity of social media and the advent of the information age, there are more and more ways for people to share their opinions and feelings online, among which Weibo is a widely used platform [4]. Weibo, as an important representative of Chinese social media, has hundreds of millions of active users and generates a large amount of text data every day [5]. These data contain users' emotional attitudes and opinions on various themes, events and products [6]. Therefore, the emotional analysis of Chinese Weibo has important practical significance and application value.

This paper will make a comprehensive and in-depth discussion on the development of Chinese Weibo sentiment analysis technology, aiming at providing a systematic perspective to understand and solve these challenges. Starting with the definition of emotional analysis, this paper expounds the important significance and application scenarios of emotional analysis in real life. At the same time, it briefly introduces the development of sentiment analysis technology, from the early dictionary matching method to the modern deep learning model, and the application of sentiment analysis technology in various fields. On the one hand, we can better understand and deal with the influence of online public opinion on social life, provide more accurate market analysis and user feedback for the government and enterprises, and make more scientific and reasonable decisions; On the other hand, through the in-depth discussion of Chinese Weibo sentiment analysis technology, it can also promote the development and progress of natural language processing and machine learning and other related fields, and lay a more solid foundation for the wide application of artificial intelligence.

2. The Present Situation and Development Trend of Sentiment Analysis Technology in Chinese Weibo

2.1. Status

With the rapid development of the Internet, especially the popularity of social media, a large number of user-generated content is generated and shared, such as blogs, forums, Weibo and so on [7]. The text information on these platforms contains users' emotions and opinions on various topics. Therefore, the emotional analysis technology, especially the Chinese Weibo emotional analysis technology, is playing an increasingly important role.

The first step of Chinese Weibo sentiment analysis is data preprocessing, which mainly includes data cleaning, word segmentation, part-of-speech tagging and stop words removal. Data cleaning mainly removes irrelevant information, such as advertisements and non-Chinese content [8]. Word segmentation and part-of-speech tagging are the steps of fine-grained processing of the text, so as to better understand and analyze the vocabulary and grammatical structure in the text. Feature extraction is a key step in the emotional analysis of Chinese Weibo. By extracting useful information from the text, such as words, phrases, sentences, etc., it is transformed into numerical features that can be understood by the computer, such as TF-IDF values, word vectors, etc. Among them, TF-IDF is a commonly used feature extraction method, which can reflect the importance of words in the text [9]. Word vector is a method of transforming words into numerical representation, which can capture the semantic information between words. In the emotional analysis of Chinese Weibo, the commonly used models and algorithms include machine learning and deep learning models. Among them, machine learning models include naive Bayes, support vector machines, logistic regression, etc. These models are widely used in sentiment analysis. Deep learning models include convolutional neural networks, cyclic neural networks, etc. These models can automatically capture the feature information in the text, thus achieving more accurate analysis.

2.2. Development Trends

With the rapid development of Internet and social media, Chinese Weibo sentiment analysis technology is becoming more and more important. This technology can be used to analyze massive user-generated content to understand the public's emotional attitudes and opinions on various themes, events and products [10]. This makes sentiment analysis technology widely used in many fields, such as market research, public opinion monitoring, decision-making and so on.

(1) Further application of deep learning model

In recent years, deep learning model has made remarkable achievements in many natural language processing tasks, including sentiment analysis. Emotional analysis of Chinese Weibo faces some challenges, such as language diversity and emotional complexity, and deep learning model has the ability to automatically capture text features, which can better deal with these problems. Therefore, it can be predicted that the deep learning model will be more and more widely used in the emotional analysis of Chinese Weibo.

(2) Fusion of multi-source data

Chinese Weibo sentiment analysis technology will not only rely on single text data, but also integrate other types of data, such as images, audio and video. These data can provide more contextual information and help to understand users' emotions and opinions more accurately. In addition, the fusion of multi-source data can also improve the accuracy and reliability of sentiment analysis.

(3) Personalization of affective computing

In the future development, Chinese Weibo sentiment analysis technology will pay more attention to the personalized needs of users. For example, personalized emotional analysis can be carried out according to users' language habits, hobbies and other information to better meet users' needs.

(4) Concerns about privacy and ethical issues

With the wide application of sentiment analysis technology, privacy and ethical issues will also receive more attention. For example, how to carry out effective emotional analysis while protecting users' privacy will be a problem that needs to be concerned and solved in the future development.

To sum up, the development trend of Chinese Weibo sentiment analysis technology will be the further application of deep learning model, the fusion of multi-source data, the personalization of sentiment calculation and the concern of privacy and ethics. This will make the Chinese Weibo sentiment analysis technology more mature and practical, and be applied in more fields.

3. Sentiment Analysis Technology under the Background of Big Data

Sentiment analysis is a natural language processing technology, which aims to automatically identify and understand the emotions and opinions expressed in the text. Sentiment analysis technology first appeared in the 1990s, when it was mainly based on rules and dictionaries. With the continuous development of machine learning and deep learning technology, sentiment analysis technology has also been widely used. At present, sentiment analysis technology has been applied in many fields, such as market research, public opinion monitoring and decision-making. Under the background of big data, Chinese Weibo sentiment analysis technology is faced with massive, high-dimensional and heterogeneous data processing and calculation problems, which need effective processing and analysis with the help of sentiment analysis technology. The sentiment analysis of Chinese Weibo under the background of big data is shown in Figure 1.

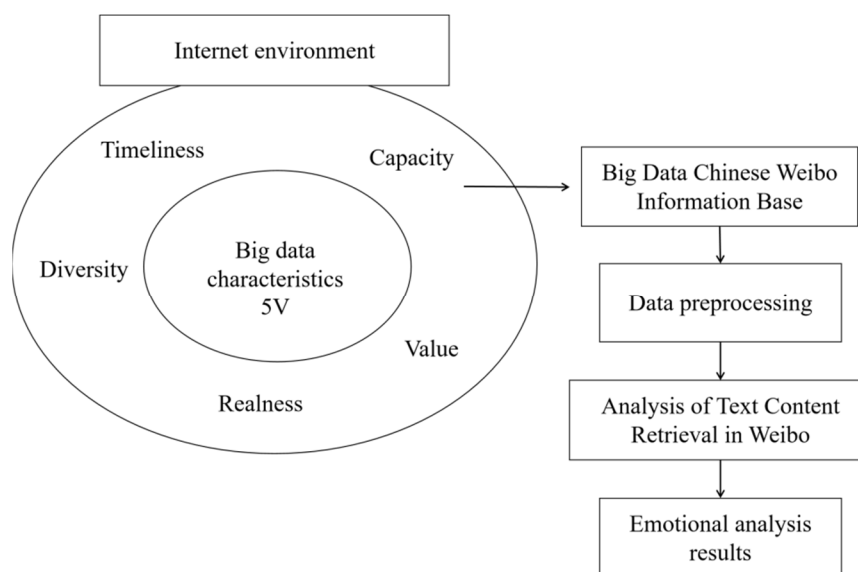


Figure 1. Emotional analysis of Chinese Weibo under the background of big data

Big data technology is widely used in Chinese Weibo sentiment analysis, mainly in the following aspects: data collection and storage: through big data technology, a large amount of data from Weibo platform and other social media platforms can be quickly collected and stored in distributed databases for subsequent processing and analysis. Data cleaning and integration: Through big data technology, collected data can be cleaned and integrated, irrelevant information can be removed, missing values can be filled in, and noise can be eliminated, so as to improve the quality and reliability of data. Feature extraction and model construction: Through big data technology, useful features can be extracted from data, and these features can be used to build emotional analysis models, such as naive Bayesian classifier, support vector machine, neural network and so on. Emotion analysis and application: Through big data technology, Weibo's text can be analyzed and applied, such as emotion dictionary construction, emotion value calculation, text classification, emotion clustering and other operations, so as to understand the public's emotional attitudes and opinions on various themes, events and products. At the same time, big data technology can also be combined with natural language

processing technology and image processing technology to achieve more comprehensive emotional analysis and application.

4. Challenges and Solutions of Emotion Analysis Technology in Chinese Weibo

4.1. Challenge

Chinese Weibo sentiment analysis technology has a broad application prospect under the background of big data, but it also faces many challenges. These challenges include, but are not limited to, dealing with massive data, dealing with unstructured data, the complexity of emotional understanding, and the diversity of languages. The details are as follows:

(1) Processing massive data

Emotion analysis of Chinese Weibo needs to deal with massive data, including text, pictures, videos and other types of data. These data are not only huge in quantity, but also complex in structure, which is difficult to be processed by traditional data processing methods. This requires the use of big data technologies, such as distributed storage and computing, data compression and other technologies, for efficient data processing and analysis.

(2) Processing unstructured data

In the emotional analysis of Chinese Weibo, unstructured data accounts for a large proportion. These data have no fixed format and structure and are difficult to be processed by traditional data processing methods. Therefore, it is necessary to study how to effectively process and analyze unstructured data by using natural language processing and machine learning technologies.

(3) The complexity of emotional understanding

The core of emotional analysis is emotional understanding, that is, understanding the emotions and opinions expressed in the text. However, emotional understanding is a very complicated problem, because the same sentence may express different emotions in different situations. This requires in-depth analysis and understanding of the text by using natural language processing, semantic understanding and deep learning technologies. The process of clustering Chinese Weibo texts is shown in Figure 2.

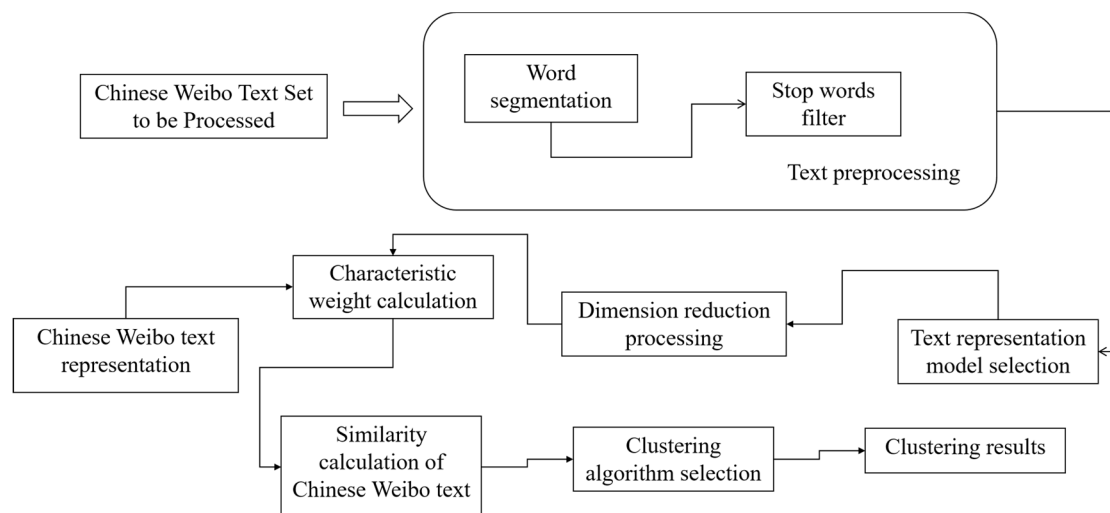


Figure 2. The process of clustering Chinese Weibo texts

(4) The diversity of languages

Emotion analysis of Chinese Weibo needs to deal with many kinds of languages, including Mandarin, dialects and online buzzwords. These languages have different expressions and characteristics, and it is difficult to deal with them in a single way. This requires the effective processing and analysis of different languages by using natural language processing and deep learning technologies.

4.2. Solution

Here are some possible solutions to the challenges mentioned in the previous section:

- (1) Processing massive data: Distributed storage and calculation methods, such as Hadoop or Spark, can be used to realize the processing and analysis of large-scale data. In addition, data compression and preprocessing technology can also be used to reduce the amount of data and improve processing efficiency.
- (2) Processing unstructured data: Natural language processing techniques, such as text segmentation, part-of-speech tagging and named entity recognition, can be used to preprocess and extract features from text data. In addition, deep learning technology can also be used to automatically learn and extract features and patterns from texts.
- (3) Complexity of emotional understanding: We can combine deep learning and natural language processing technology, and use word vectors, sentence vectors and other representation methods to classify or predict emotional values of texts. In addition, transfer learning and multi-task learning can be used to improve the generalization performance and robustness of the model.
- (4) Language diversity: Special corpora and dictionaries can be established for different language types and expressions, and natural language processing technology can be used for text segmentation, part-of-speech tagging, named entity recognition and other operations. In addition, speech recognition technology can be used to transcribe speech into text for emotional analysis.

It should be noted that in the emotional analysis of Chinese Weibo, it is necessary to protect users' privacy and data security, and at the same time, it is also necessary to pay attention to the handling of ethical issues. For example, the results of emotional analysis may affect the public's freedom of speech and right of expression, so it is necessary to fully consider various factors in application and formulate corresponding norms and standards to ensure fairness and objectivity.

5. Conclusion

Chinese Weibo sentiment analysis technology is one of the important research directions in the field of natural language processing in the era of big data, which has a wide application prospect in various fields. However, the development of sentiment analysis technology in Chinese Weibo faces many challenges, such as massive data processing, unstructured data processing, the complexity of sentiment understanding and the diversity of languages. In view of these challenges, this paper gives the corresponding solutions. Through the research of Chinese Weibo sentiment analysis technology, we can better understand and deal with the influence of online public opinion on social life, provide more accurate market analysis and user feedback for the government and enterprises, and make more scientific and reasonable decisions. This research not only has important practical significance for the development of Chinese Weibo sentiment analysis technology, but also has certain academic value and practical guiding significance for promoting the development of related fields. It is expected that the discussion in this paper will arouse more readers' attention and thinking on the emotional analysis technology of Chinese Weibo, and promote the further development and progress in this field. At the same time, I also believe that with the continuous progress of technology and the increasing demand for application, the Chinese Weibo sentiment analysis technology will be widely used and developed in more fields.

References

- [1] Bao Shuhua, Shi Yingxin. Emotional analysis of domestic film reviews under the condition of big data [J]. journal of hulunbeier college, 2022(002):030.
- [2] Lu Yanxia, Wu Di, Huang Chuanlin. Emotional analysis of Weibo users facing public emergencies under big data environment [J]. Software Engineer, 2017, 020(001):45-48.
- [3] Liu Ping. The public's emotional tendentiousness towards domestic football players-Weibo's emotional analysis based on big data [J]. Sports Science and Technology Literature Bulletin, 2022, 30(3):246-250.

- [4] Shi Zhenhua. Cognitive and Emotional Analysis of Chengdu Night Tourists Based on Weibo Big Data [J]. Journal of Sichuan Cadre Correspondence College, 2021, 000(003):82-89.
- [5] Li Tongtong, Guo Xuning, Zhou Yanli, et al. Analysis of the emotional state of college students under the COVID-19 epidemic and its influencing factors-evidence based on Weibo's text mining [J]. Open Learning Research, 2022, 27(5):10.
- [6] Zhou Yuan. A method for emotional analysis of social big data [J]. China Strategic Emerging Industries, 2019, (010):194.
- [7] Li Zekui, Li Xueting, Zhao Yanyan. Analysis of the reasons for the emotional distribution of hot events in Weibo [J]. journal of chinese information, 2018, 32(1):8.
- [8] Cheng Ziwen, Zeng Hao, Jiang Bin, et al. Analysis and design of Weibo public opinion text relevance under the background of big data [J]. Modern Information Technology, 2020(018):004.
- [9] Wang Xiwei, Xing Yunfei, Wei Yanan, et al. Research on the construction of big data-driven emotional theme classification model of public opinion users in social networks-taking the theme of "immigration" as an example [J]. 2021(2020-1):29-38.
- [10] Liu Qihui, Chai Yumei, Liu Zhen. Chinese Weibo Emotion Analysis Model SR-CBOW[J]. Miniature Microcomputer System, 2018, 39(008):1693-1699.