Proceedings of the 2023 International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/36/20230440

Text emotional analysis in Natural Language Processing

Shangyue Lin

Hangzhou No.2 High School of Zhe Jiang Province, Hangzhou, Zhejiang, 310059, China

220712031@njnu.edu.cn

Abstract. Natural Language Processing (NLP) is an important research direction in the field of artificial intelligence, aimed at enabling computers to understand and process human natural language. Emotional analysis refers to the use of computer technology to automatically recognize and classify the emotional tendencies expressed in text, such as positive, negative, or neutral. This technology can be applied to fields such as social media analysis, public opinion monitoring, market research, etc. The goal is to enable computers to understand and interpret emotional information in text like humans. Text sentiment analysis, as a key task of NLP, aims to identify and classify emotional tendencies expressed in text. This article will delve into the field of text sentiment analysis in natural language processing, focusing on the role and application of standard sentiment dictionaries and corpora in sentiment analysis. Reviewing relevant research conclusions, providing reference and inspiration for achieving more accurate and comprehensive text sentiment analysis through research and exploration.

Keywords: Natural Language Processing, Artificial Intelligence, Emotional Analysis.

1. Introduction

Text sentiment analysis is an important task in natural language processing, which studies the interaction between human language and computers. Text sentiment analysis focuses on identifying and understanding emotions or emotions in text. Natural language processing includes multiple tasks and techniques, among which text sentiment analysis is one of them [1]. By applying natural language processing technology, emotional content in text can be analyzed and understood, thereby extracting emotional information [2]. Text sentiment analysis uses various methods and techniques, including rule-based methods, dictionary matching, machine learning, and deep learning. These methods can help classify text into different sentiment categories (such as positive, negative, and neutral), evaluate sentiment intensity, and identify targets or entities related to emotions [3]. Text sentiment analysis has a wide range of applications in various fields. It can be used for social media analysis, brand reputation management, public opinion monitoring, market research, user comment analysis, and more. By analyzing the emotional aspects of the text, the public's attitudes and emotions towards specific topics, products, or services can be obtained, assisting in decision-making and providing insights [4].Therefore, text sentiment analysis is an important task in the field of natural language processing, which utilizes natural language processing technology to solve the problem of emotional content analysis and understanding in texts [5].In recent years, with the significant increase in social media, online comments, and user feedback, it has become increasingly important to obtain people's emotional tendencies

DOI: 10.54254/2755-2721/36/20230440

towards products, services, events, or topics. Text sentiment analysis can automatically classify, intensity analyze, and target recognize emotions in text by utilizing natural language processing techniques, thereby helping us gain insight into public attitudes and emotions. In terms of text sentiment classification, traditional machine learning methods such as support vector machines (SVM) and naive Bayesian classifiers are often used, which typically rely on feature extraction and selection, including word bag models, n-gram features, and part of speech tagging [6].In recent years, deep learning methods such as Convolutional Neural Networks (CNN) and Long Short Term Memory Networks (LSTM) have also made significant progress, being able to automatically learn feature representations and perform well in emotion classification tasks [7].

Emotional intensity analysis aims to determine the degree or intensity of emotions in a text. This task can be achieved through dictionary matching methods, rule-based methods, and machine learning algorithms. An emotion dictionary is a commonly used resource that contains information about the correlation between words and phrases and emotional polarity or intensity [8]. By using dictionary matching or rule-based methods, the number and weight of emotional words in the text can be calculated, and the emotional intensity can be estimated based on this. Emotional target recognition refers to identifying entities or objects related to emotions in text. This task is usually achieved by combining named entity recognition technology and sentiment classification methods [9]. By annotating entities and corresponding emotional categories in the text, it is possible to gain a deeper understanding of the relationship between emotions and specific goals, thereby providing more detailed and accurate emotional analysis results [10]. Although there has been some progress in text sentiment analysis, it still faces some challenges. Semantic ambiguity, subjectivity of emotional expression, and noise and incomplete information in the text are all issues that need to be addressed. To address these challenges, researchers are constantly exploring new methods and technologies, such as expanding emotional dictionaries, transfer learning, and attention mechanisms [11]. By utilizing various technologies and methods, it is possible to better understand and analyze the emotional content in texts, providing valuable insights and decision-making support for social media analysis, public opinion monitoring, brand reputation management, and other fields. With the continuous development of technology, it is believed that text sentiment analysis will achieve more accurate and practical results in the future.

2. Method

NLP is an important branch of computer science dedicated to understanding and processing human language. One of the key tasks is text sentiment analysis, which infers emotional states by analyzing emotional tendencies in the text. In the past few years, CNN methods have achieved significant results in text sentiment analysis. CNN is a deep learning model widely used in the field of image recognition, which has been introduced into text processing in recent years. It captures contextual information by extracting local features from input data, which is crucial for sentiment analysis. The core components of CNN include convolutional layer, pooling layer, and fully connected layer. Convolutional layers scan text sequences through sliding windows to extract features of different sizes. The pooling layer preserves the most important features through downsampling operations [12]. Finally, the fully connected layer maps the extracted features to the emotional label space. The process of sentiment dictionary analysis is shown in Figure 1.

DOI: 10.54254/2755-2721/36/20230440

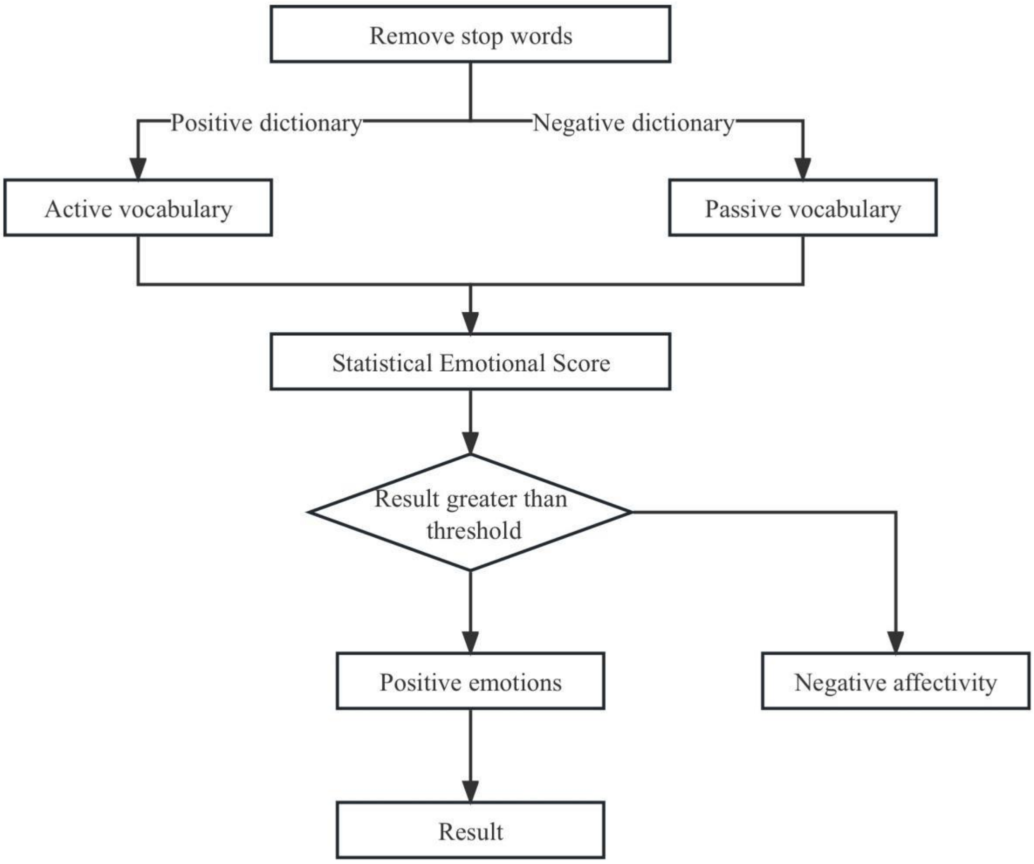

Figure 1. Emotional Dictionary Analysis Process.

The first step in text sentiment analysis using CNN is to convert the text into a numerical representation. Common methods include the bag-of-words model, word embedding, and character-level representation. The lexical bag model represents text as a fixed-length vector, where each dimension corresponds to a lexical item and the frequency of occurrence of each lexical item is calculated. Word embedding is a technology for mapping words to continuous vector spaces, which captures the semantic relationships between words. Character level representation transforms text into character sequences, and uses convolutional neural networks to extract features directly from the character level [13]. In CNN, it is important to use convolution kernels for feature extraction. Convolution kernels can capture local features of varying lengths, such as phrases, sentence structures, and emotional expressions. Through multiple convolution kernels of different sizes, CNN can simultaneously learn feature representations at different levels. In addition, regularization techniques such as dropout and L2 regularization are often used to enhance the generalization ability of the model.

When training a CNN model, it is necessary to label good emotional labels as supervised signals. A common method is to use supervised learning algorithms such as multi-classifiers or binary classifiers [14]. In addition, external resources such as sentiment dictionaries can be used to enhance the performance of the model. It is worth noting that data preprocessing and feature selection are also crucial in text sentiment analysis, and models can be optimized through methods such as removing stop words, stem extraction, and punctuation processing. Although CNN performs well in text sentiment analysis, there are still some challenges. First, CNN may lose important contextual information when processing long texts. Secondly, over-reliance on the extraction of local features may lead to a high sensitivity of the model to input data. Additionally, for sentiment analysis in non-English languages and specific domains, the generalization ability of CNN models may be limited [15]. CNN is a powerful text sentiment analysis method that captures context information in the text through extracting local features and has excellent classification performance. However, further research is needed to address the challenges and potential solutions for CNN models when dealing with long texts and non-English text. For long texts, layered convolution structures can be used, where the input text is divided into multiple sub-sequences and context information is captured over a longer range by stacking multiple convolutional layers. Additionally, attention mechanisms can be used to dynamically focus on important sections related to emotions, thus improving the comprehension of long text [16].

A common challenge for non English texts is the differences in expression and vocabulary between different languages. One way to solve this problem is to introduce multilingual word embedding, which maps the word vector space of multiple languages to a shared semantic space. This can establish connections between different languages, improve the generalization ability of the model, and reduce data scarcity issues. In addition to the limitations and challenges mentioned above, there are also other factors that need to be considered, such as the size of the model, hyperparameter adjustment, and the quality and quantity of training data. In addition, the selection of evaluation indicators is also an important issue. Common evaluation indicators include accuracy, precision, recall, and F1 value, but in sentiment analysis, emotional orientation measures such as positivity, negativity, and neutrality can also be considered [17]. In summary, convolutional neural networks are an effective method for text sentiment analysis and have broad application prospects. Through appropriate model design and data processing techniques, the challenges of CNN processing long and non English texts can be overcome, and the accuracy and robustness of sentiment analysis tasks can be improved. In the future, more complex network structures and more advanced pre training models can be further explored to promote the development of text sentiment analysis [18].

3. Results

3.1. Extraction of emotional words

Emotional analysis is an important task in the field of natural language processing, aimed at identifying and understanding the emotions and emotional tendencies expressed in text. Extracting emotional words is a key measure in sentiment analysis. By identifying emotional words in text, can better understand and analyze the emotions in the text. The extraction of emotional words is shown in Figure 2.

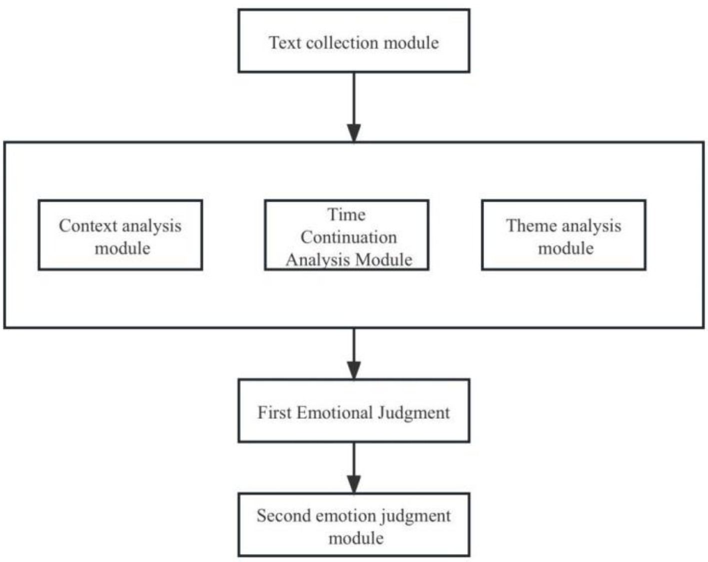

Figure 2. Extraction of Emotional Words.

DOI: 10.54254/2755-2721/36/20230440

An emotional dictionary is a vocabulary that includes words and their corresponding emotional polarity. Based on this vocabulary, words in the text can be matched with emotional polarity to identify emotional words. For example, words with positive emotions such as "happy" and "like", and words with negative emotions such as "sad" and "dislike". By calculating the number of positive and negative emotion words in the text, the emotional tendency of the text can be obtained. In addition to using a dictionary, machine learning methods can also be used to extract emotional words, which typically require a training dataset annotated with emotions as input to learn the features and patterns of emotional words during the training process. Commonly used machine learning algorithms include naive Bayes, support vector machines, and deep learning methods. By using machine learning methods, text emotions can be predicted based on features such as word frequency and part of speech [19]. The rule-based method is a way to directly define the extraction rules of emotional words, designed based on grammar, part of speech tagging, phrase, and other aspects. For example, by defining adjectives that appear after noun phrases as emotional words, words with emotional tendencies can be extracted. In recent years, with the development of deep learning technology, emotion analysis methods based on deep learning have gradually become popular. Deep learning models, such as recurrent neural networks (RNN), short-term memory networks (LSTM), and transformers, can capture contextual information in text, thereby improving the accuracy of emotion word recognition [20]. Overall, emotion word extraction is an important step in emotion analysis. By using dictionary methods, machine learning methods, rule-based methods, and deep learning methods, can effectively extract emotional words from text and assist in sentiment analysis. Different methods are suitable for different scenarios and tasks, and in practical applications, appropriate methods need to be selected based on specific situations to extract and analyze emotional words [21].

3.2. Construction of an emotional dictionary

Building an emotion dictionary is an important step in text sentiment analysis. An emotion dictionary is a vocabulary that includes words and their corresponding emotional polarity (such as positive and negative), used to identify emotional tendencies in text. Firstly, it is necessary to collect a set of seed words that have already been labeled with emotional polarity. These seed words can be obtained from existing emotional dictionaries, manually annotated datasets, or professional knowledge, and should cover multiple fields and different emotional polarities [22]. Using seed vocabulary as a starting point, the emotional vocabulary list can be expanded through various methods. Using a synonym dictionary or word embedding model, add the synonyms of the seed vocabulary to the emotional vocabulary list. Add words with opposite emotional polarity to the emotional vocabulary list as seed words. Generate derived words related to seed vocabulary by adding prefixes, suffixes, or changing word form. Collect and expand emotional vocabulary related to a specific field [23]. For the expanded vocabulary list, it is necessary to annotate emotional polarity. This can be achieved through manual annotation, crowdsourcing platforms, or rule-based methods. Markers need to assign corresponding emotional polarity labels to each vocabulary based on their own understanding and knowledge, such as positive, negative, or neutral.

After annotating emotional polarity, it is necessary to filter and filter the vocabulary list. The purpose of this step is to eliminate irrelevant or inaccurate vocabulary and improve the quality of emotional dictionaries. Filter according to the following aspects: (1) Semantic consistency: Check whether the semantics of emotional words are consistent with their emotional polarity, and avoid words with semantic ambiguity. (2) Context Dependency: Considering that the emotional expressions of vocabulary may differ in different contexts, emotional words can be adjusted or filtered based on contextual information. (3) Domain adaptability: If emotional vocabulary is constructed for a specific domain, further screening can be carried out based on the characteristics of the domain [24]. The construction of an emotional dictionary is an iterative process. When using an emotion dictionary for sentiment analysis, the output results of the model can be evaluated, and the emotion dictionary can be updated and optimized based on feedback information, which involves adding new vocabulary, modifying or deleting existing vocabulary to adapt to constantly changing language use and emotional expression.

DOI: 10.54254/2755-2721/36/20230440

The construction of an emotional dictionary is a complex task that is influenced by factors such as ambiguity, cultural differences, and subjectivity. Therefore, building a high-quality emotional dictionary requires comprehensive consideration of multiple key points:

First, multi-source criteria need to be met. In addition to seed vocabulary, emotional vocabulary can also be obtained from various sources, such as domain expert knowledge, social media data, online reviews, etc. Secondly, manual annotation and automatic mining are required. The construction of an emotion dictionary can be combined with manual annotation and automatic mining methods. Manual annotation can provide high-precision emotional polarity labels, while automatic mining can quickly expand the emotional vocabulary list. Third, preprocessing and standardization are required. Before constructing emotional vocabulary, it is necessary to preprocess the text, such as removing stop words, stemming, etc. In addition, it is necessary to standardize emotional words to ensure consistency and accuracy. Fourthly, the domain adaptability needs to be looked at. For sentiment analysis in specific fields, the adaptability and accuracy of sentiment dictionaries can be improved by introducing domain related corpus or features. Fifth, pay attention to the need to constantly update maintenance. The construction of emotional vocabulary is an ongoing process, and with changes in language and society, it is necessary to regularly update and maintain emotional dictionaries to maintain their effectiveness and reliability.

3.3. Application of Natural Language Processing

Collect text data containing emotional information from social media, news articles, online comments, and other fields. In the data preprocessing stage, common text preprocessing operations such as removing special characters, converting to lowercase, word segmentation, and removing stop words can be performed to prepare data for emotional analysis. An emotional dictionary contains words and their corresponding emotional polarity. Building an emotional dictionary is an important step in emotional analysis, using existing emotional dictionaries such as SentiWordNet and AFINN to determine the emotional words in the text. Using emotion dictionaries or other methods to extract emotional words from text, that is, words with emotional tendencies. In addition, other feature representation methods such as word frequency, part of speech tagging, syntactic analysis, etc. can also be used to capture the emotional features of the text [25].

Based on the extracted emotional words and feature representations, different algorithms or rules are used to calculate the emotional polarity score of the text, calculate the proportion or weighted score of positive and negative emotional words, and determine the overall emotional tendency. For more complex emotion analysis tasks, machine learning or deep learning models such as feature engineering, selecting appropriate algorithms or model architectures can be used for training and prediction, and emotion annotated training data can be used for model training and evaluation. Based on the results of sentiment analysis, explain the emotional tendencies of the text and apply them to various application scenarios such as public opinion analysis, product review analysis, and social media sentiment monitoring. According to requirements, further visualization or information aggregation of the results can be carried out to support decision-making and insight discovery.

4. Application of research results

4.1. Basic problem analysis

Emotional classification is the classification of text into different emotional categories, usually including positive, negative, and neutral, aimed at determining the overall emotional tendency expressed in the text. The methods for solving emotion classification problems can be based on machine learning, deep learning, or rule matching. Common techniques include using feature engineering to extract text features, and using classification algorithms such as naive Bayes, support vector machines, deep neural networks, etc. for training and prediction.

Emotional intensity analysis focuses on the degree or intensity of emotions expressed in the text, and usually requires evaluating the emotional words in the text and calculating the intensity score of emotions. A common method is to extract sentiment words through sentiment dictionaries or pre trained sentiment models and perform weighted calculations, which is of great significance in some application scenarios such as sentiment intensity assessment in public opinion analysis and sentiment polarity rating in product reviews. Emotional target recognition aims to determine the emotional objects or targets involved in the text. The text usually mentions multiple entities or viewpoints, while the emotional target recognition task is to identify the specific targets targeted by the text and judge the emotions expressed. The method to solve this problem can be based on technologies such as named entity recognition, keyword matching, and semantic role annotation. Emotional target recognition has important application value in scenarios such as social media analysis, public opinion monitoring, and product review analysis.

4.2. Emotional analysis of new social media

For emotional analysis of new social media, data needs to be obtained from corresponding social media platforms such as Twitter, Weibo, Instagram, etc. When preprocessing data, appropriate adjustments need to be made to the characteristics of social media text, such as processing URL links, removing @ user tags, processing emoticons and special symbols, etc. Emotional dictionaries are an important resource for sentiment analysis. For sentiment analysis on new social media, existing sentiment dictionaries can be used and appropriately expanded and modified in conjunction with social media contexts. In addition, user generated tags or comment data can be used to expand the sentiment dictionary to better adapt to the language characteristics of social media. In addition to emotional vocabulary, the effectiveness of emotional analysis can also be enhanced by extracting other text features. For example, methods such as word bag models, n-gram features, and part of speech tagging can be used to capture the emotional features of text. In addition, metadata information such as symbolic emoticons, hash tags, and user emotional tendencies can also be considered.

For emotional analysis of new social media, machine learning algorithms can be used for modeling and prediction. Common methods include naive Bayesian classifiers, support vector machines, random forests, and deep learning models (such as convolutional neural networks, long-term and short-term memory networks, etc.). These methods need to be combined with annotated training data for model training, and optimized using techniques such as feature engineering and parameter tuning. When conducting emotional analysis on new social media, it is necessary to consider the contextual information of social media texts. Text on social media platforms is usually short and contains a large number of abbreviations, typos, and slang. Therefore, it is necessary to consider technical means of contextual completion and correction to more accurately understand emotional tendencies. In recent years, deep learning methods have made significant progress in sentiment analysis. For example, using word embedding techniques such as Word2Vec and GloVe to learn distributed representations of words, and using recurrent neural networks such as long short-term memory networks to model sentence or document level emotions, these methods typically better handle semantic information and contextual relationships in social media texts.

4.3. Application oriented emotional analysis

NLP's text sentiment analysis plays an important role in various fields of application. Emotional analysis aims to identify, classify, and evaluate the emotional tendencies expressed in text, helping people understand and analyze the emotional information contained in large-scale text data.

Public opinion monitoring refers to the real-time analysis and monitoring of massive text data such as social media, news reports, and online comments through sentiment analysis technology, helping enterprises, organizations, or government agencies understand the public's emotional tendencies towards specific events, products, or brands, and timely detect and solve potential crises. Based on effective monitoring and management of public opinion, reputation management strategies can be improved, public image and word-of-mouth can be enhanced. Emotional analysis can be used to analyze consumers' evaluations and feedback on products or services, helping businesses understand their satisfaction, needs, and preferences. Analyzing product reviews, social media discussions, and user feedback can provide information about product quality, functionality, and user experience. Accurate emotional analysis can be applied to social media analysis, such as analyzing user emotional tendencies on platforms such as Twitter and Weibo, tracking topic trends, and inferring user emotions, which helps to understand public opinion trends, user preferences, and behavioral patterns. In addition, in recommendation systems, sentiment analysis can be used to personalized recommend content, matching corresponding products, music, movies, etc. based on users' emotional preferences. In addition, it can also be applied to building emotional interaction systems and emotional robots, analyzing users' emotional states and feedback, achieving a more intelligent and personalized user experience, services, and support, suitable for fields such as voice assistants, intelligent customer service, and virtual emotional partners. Assist marketers in evaluating the effectiveness of advertising, promotional activities, and marketing strategies, analyzing users' emotional reactions to advertisements or brands, and adjusting and optimizing advertising content, positioning, and communication channels to improve the effectiveness and return on advertising placement. In addition, emotional analysis can also help understand consumers' views on brand image and market competition strategies. Help educators understand students' emotional states and needs, and provide personalized guidance and support based on emotional analysis results. In the field of mental health, emotional analysis can be used for automated emotional monitoring and diagnosis, helping with psychological counseling and treatment processes.

Emotional analysis has multiple applications in the field of news media. Firstly, emotional analysis can help news organizations analyze readers' emotional reactions to news reports. By analyzing structured data from reader comments, social media discussions, and emotional analysis, news organizations can understand the public's emotional tendencies towards different news events and adjust news content and reporting perspectives accordingly. Secondly, emotional analysis can also be used for editing decisions. News editors can use sentiment analysis techniques to evaluate the emotional color of articles to ensure objectivity and balance in reporting. Emotional analysis can help editors understand the emotional information contained in texts, thereby better selecting vocabulary and sentence structures to convey the desired emotional effects. In addition, sentiment analysis can also be applied to false news detection. By analyzing news articles or content on social media, potential emotional biases, exaggerations, or misleading expressions can be identified. This technology helps to improve the accuracy and credibility of news reporting. Finally, emotional analysis can be applied to political analysis. By analyzing the emotional attitudes and expressions of voters, one can gain insight into the public's attitude and level of support towards political candidates, political parties, or specific policies. This is of great significance for shaping the image of candidates, formulating campaign strategies, and guiding public opinion. In summary, text sentiment analysis in natural language processing has been widely applied in various fields. From public opinion monitoring to marketing, from education to mental health, from news media to political analysis, emotional analysis provides people with deeper insights and decision-making support. With the continuous development and innovation of technology, emotional analysis will further promote the intelligent and personalized development of various industries.

4.4. Standard emotional dictionary and corpus

In text sentiment analysis in natural language processing, standard sentiment dictionaries and corpora are two key components, providing vocabulary resources and training data for sentiment analysis, which can help researchers and developers identify and classify emotional tendencies. A standard emotional dictionary is a dictionary resource that contains emotional vocabulary and its corresponding emotional polarity, dividing vocabulary into positive, negative, or neutral emotions, and assigning emotional scores to each vocabulary. Emotional dictionaries are usually manually annotated by experts or constructed through machine learning methods, and typically include adjectives, adverbs, verbs, and other parts of speech to describe things, behaviors, and emotional states. A standard sentiment dictionary can serve as a basic resource for sentiment analysis tasks, for algorithms and models to use. Based on the emotional vocabulary appearing in the text, the overall emotional tendency of the text expression can be estimated by calculating the score and weighted sum of the emotional vocabulary.

The method based on sentiment dictionaries is simple and intuitive, but requires accurate emotional vocabulary resources. A corpus is a collection of large-scale text data used to train and evaluate sentiment analysis models. A corpus includes text collected from various sources such as social media posts, news articles, comments, etc. These texts are typically labeled as positive, negative, or neutral emotions to provide training data. A large-scale corpus can provide more comprehensive text coverage and emotional expression, which helps improve the generalization ability and accuracy of the model. In addition, a diverse corpus can cover texts from different fields, genres, and cultural backgrounds, enabling the model to have a wider range of application capabilities.

By using a corpus for training, the sentiment analysis model can learn the mapping relationship from text features to emotional tendencies, and use machine learning techniques such as supervised learning and deep learning to associate text features in the corpus with their emotional categories. Standard emotion dictionaries and corpora need to be customized or adapted according to specific tasks and languages, and the expression of emotions varies in different fields, languages, and cultures. Therefore, it is necessary to customize dictionaries and corpora according to specific needs. Standard sentiment dictionaries and corpora are important resources for text sentiment analysis in natural language processing, providing vocabulary resources and training data to help identify and classify sentiment tendencies in text, providing a foundation for sentiment analysis research and application, and promoting the development of emotional intelligence technology.

5. Conclusion

Text sentiment analysis in natural language processing is a widely studied field, and research has shown that emotional vocabulary plays a crucial role in sentiment analysis. Identifying and utilizing emotional vocabulary can help determine the emotional tendencies of texts. Therefore, constructing accurate and comprehensive emotional vocabulary resources is crucial for the success of sentiment analysis. Context information has a significant impact on the effectiveness of sentiment analysis. The same vocabulary expresses different emotional meanings in different contexts, and considering contextual information such as word order, sentence structure, and semantic relationships can improve the accuracy of sentiment analysis. Traditional machine learning methods such as Support Vector Machines (SVM), Decision Trees, and Naive Bayes have achieved some success in sentiment analysis. With the development of deep learning technology, deep neural network models such as Convolutional Neural Networks (CNN), Recursive Neural Networks (RNNs), and Transformer models have achieved better performance in sentiment analysis tasks. Emotional analysis faces an important challenge, which is to classify emotions for texts in different fields. Due to significant differences in language usage, vocabulary, and emotional expression in different fields, universal emotional analysis models often cannot be well generalized. In future research, propose methods for domain adaptation and transfer learning, utilizing source domain data or annotated data to transfer the emotional analysis ability of the model to the target domain, in order to improve the performance of cross domain emotional analysis.

References

[1] Nijhawan T, Attigeri G, Ananthakrishna T. 2022 Stress detection using natural language processing and machine learning over social interactions. J. Big D. 9(1): 1-24

[2] Tanana M J, Soma C S, Kuo P B, et al. 2021 How do you feel? Using natural language processing to automatically rate emotion in psychotherapy[J]. Behav. Res. Methods (13):1-14

[3] Maulud D H, Zeebaree S R M, Jacksi K, et al. 2021 State of art for semantic analysis of natural language processing[J]. QAJ 1(2): 21-28

[4] Darwish K, Habash N, Abbas M, et al. 2021 A panoramic survey of natural language processing in the Arab world[J]. CMR 64(4): 72-81

[5] Wignell P, Chai K, Tan S, et al. 2021 Natural language understanding and multimodal discourse analysis for interpreting extremist communications and the re-use of these materials online[J]. TPV 33(1): 71-95

[6] Zhang T, Schoene A M, Ji S, et al. 2022 Natural language processing applied to mental illness detection: a narrative review[J]. NPJ Digit Med 5(1): 46
[7] Peng S, Cao L, Zhou Y, et al. 2022 A survey on deep learning for textual emotion analysis in social networks[J]. DICON 8(5): 745-762
[8] Khyani D, Siddhartha B S, Niveditha N M, et al. 2021 An interpretation of lemmatization and stemming in natural language processing[J]. J. Uni. Shanghai. Sci. Tech. 22(10): 350-357
[9] Murthy A R, Kumar K M A. 2021 A review of different approaches for detecting emotion from text[C]//IOP Conference Series: Materials Science and Engineering. IOP Publishing, 1110(1): 012009
[10] Torregrosa J, Bello-Orgaz G, Martínez-Cámara E, et al. 2023 A survey on extremism analysis using natural language processing: definitions, literature review, trends and challenges[J]. J. Ambient Intell. Hum. Comput. 14(8): 9869-9905
[11] Zad S, Heidari M, James Jr H, et al. 2021 Emotion detection of textual data: An interdisciplinary survey[C]//2021 IEEE World AI IoT Congress (AIIoT). IEEE, 0255-0261
[12] Fan C, Yuan C, Gui L, et al. 2021 Multi-task sequence tagging for emotion-cause pair extraction via tag distribution refinement[J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 29: 2339-2350
[13] Acheampong F A, Nunoo-Mensah H, Chen W. 2021 Transformer models for text-based emotion detection: a review of BERT-based approaches[J]. ARTIF INTELL REV (09): 1-41
[14] Chen Q, Leaman R, Allot A, et al. 2021 Artificial intelligence in action: addressing the COVID-19 pandemic with natural language processing[J]. Ann. Rev. Biomed. Da. S. (04): 313-339
[15] Levis M, Westgate C L, Gui J, et al. 2021 Natural language processing of clinical mental health notes may add predictive value to existing suicide risk models[J]. PSYCHOL MED 51(8): 1382-1391
[16] Khanbhai M, Anyadi P, Symons J, et al. 2021 Applying natural language processing and machine learning techniques to patient experience feedback: a systematic review[J]. BMJ HEALTH CARE INFO 28(1):112-113
[17] Nandwani P, Verma R. 2021 A review on sentiment analysis and emotion detection from text[J]. SNAM 11(1): 81-81
[18] Sean B. Findings of the shared task on Emotion Analysis in Tamil[C]//Proceedings of the Second Workshop on Speech and Language Technologies for Dravidian Languages. 2022,(13): 279-285
[19] Pérez J M, Furman D A, Alemany L A, et al. 2021 Robertuito: a pre-trained language model for social media text in spanish[J]. arXiv preprint arXiv.(96):87-88
[20] Choudrie J, Patil S, Kotecha K, et al. 2021 Applying and understanding an advanced, novel deep learning approach: A Covid 19, text based, emotions analysis study[J]. ISF (23): 1431-1465
[21] Nandwani P, Verma R. 2021 A review on sentiment analysis and emotion detection from text[J]. SNAM 11(1): 81
[22] Sean B. Findings of the shared task on Emotion Analysis in Tamil[C]//Proceedings of the Second Workshop on Speech and Language Technologies for Dravidian Languages. 2022(13): 279-285
[23] Graterol W, Diaz-Amado J, Cardinale Y, et al. 2021 Emotion detection for social robots based on NLP transformers and an emotion ontology[J]. Sensors, 21(4): 1322
[24] Bhowmik N R, Arifuzzaman M, Mondal M R H, et al. 2021 Bangla text sentiment analysis using supervised machine learning with extended lexicon dictionary[J]. NLPR 1(3-4): 34-45
[25] Harrison C J, Sidey-Gibbons C J. 2021 Machine learning in medicine: a practical introduction to natural language processing[J]. BMC MED RES METHODOL 21(1): 1-11