Journal of Computing and Electronic Information Management
ISSN: 2413-1660 | Vol. 15, No. 3, 2024

A Review: Text Sentiment Analysis Methods

Mengtao Zhou

School of Management, Huazhong University of Science and Technology, Wuhan, Hubei 430074, China

Abstract: This paper conducts a literature review based on the Web of Science Core Collection and CNKI databases, employing concepts, methods, and techniques related to text sentiment analysis to construct search queries. It retrieves and analyzes relevant literature on text sentiment analysis from the past decade, performing a thematic analysis to summarize and categorize the mainstream methods used in sentiment analysis, and discusses their strengths and weaknesses. The analysis identifies three primary approaches to sentiment analysis: dictionary-based, machine learning-based, and deep learning-based methods. Each method has its merits, drawbacks, and specific application scenarios. Additionally, within deep learning, self-attention mechanisms and pre-trained models have become key research areas. The paper provides an overview of sentiment analysis methods from a broad technical perspective without delving into specific details within various fields or introducing and comparing cutting-edge methods, thus presenting some limitations. Finally, it summarizes the requirements and application scenarios for the three models and offers corresponding recommendations.

Keywords: Emotion Classification; Sentiment Lexicon; Machine Learning; Deep Learning.

1. Introduction

In the current wave of the big data era, computational science is transitioning into data-intensive science, and data-driven decision-making has become central to various industries. By analyzing patterns and trends in vast amounts of data, businesses can gain insights into market dynamics and optimize their operational strategies. Furthermore, social media platforms such as Facebook, Weibo, and Twitter have become important channels for users to express their opinions and emotions. The textual data on these platforms reflects public sentiment and social hot topics. For businesses, the ability to analyze the emotional tone of this unstructured text is crucial, as it provides essential references for brand management, public opinion monitoring, market research, targeted marketing, and risk forecasting, laying a data foundation.

Sentiment analysis, also known as opinion mining, aims to extract emotion-laden subjective texts from users' social network activities (e.g., comments, blogs), analyze, process, summarize, and infer these texts, and use certain sentiment score metrics to quantify qualitative data. Researchers generally classify the sentiment polarity of texts into three granular levels: document-level, sentence-level, and aspect-level, based on the granularity of the text.[1]. The text granularity classification method has become a more commonly used classification approach in research. This paper primarily studies the methods of text sentiment analysis. It conducts a keyword analysis of the literature retrieved from the Web of Science core database and China National Knowledge Infrastructure, summarizes the main models and methods of sentiment analysis in recent years, and provides a brief comparative analysis of domestic and international literature. Subsequently, based on the existing literature, text sentiment analysis methods are classified into three types: sentiment lexicon-based, machine learning-based, and deep learning-based. Each of these has simple subcategories. This paper compares and analyzes the characteristics of these methods, addresses gaps in previous literature, focuses on introducing the most cutting-edge methods in deep learning, and finally explains the application scenarios and requirements of each method.

2. Current Status of Research on Text Sentiment Analysis

2.1. Domestic Literature Review

This paper selects the well-known and information-rich database, China National Knowledge Infrastructure (CNKI), as the data platform. Using the advanced search function, the keywords 'sentiment analysis,' 'opinion mining,' 'viewpoint mining,' 'emotion analysis,' and 'sentiment classification' were selected. Then, additional filtering was done using the terms 'algorithm,' 'model,' 'mechanism,' 'learning,' and 'lexicon,' with the time range set from January 1, 2014, to September 1, 2024, covering literature from the past decade. A secondary search was conducted within the search results, and after filtering and selecting core journals, 1,625 Chinese articles were selected. Out of these, more than 30 articles were chosen for detailed reading. The search for articles on the topic of 'sentiment analysis' was further refined by narrowing the source categories to include only Peking University Core Journals, CSSCI, and EI databases. This focused search yielded a total of 360 articles, ensuring a more targeted and relevant collection of literature for in-depth analysis. These were then analyzed using CiteSpace to identify key thematic keywords in this research field, as shown in Figure 1.

As shown in Figure 1, the primary application areas of sentiment analysis include online reviews, online public opinion, Weibo, and user comments. The primary approaches utilized include methods based on sentiment lexicons, machine learning techniques, and deep learning algorithms. In China, there is a tendency towards research focused on deep learning methods, which constitutes the main research direction.

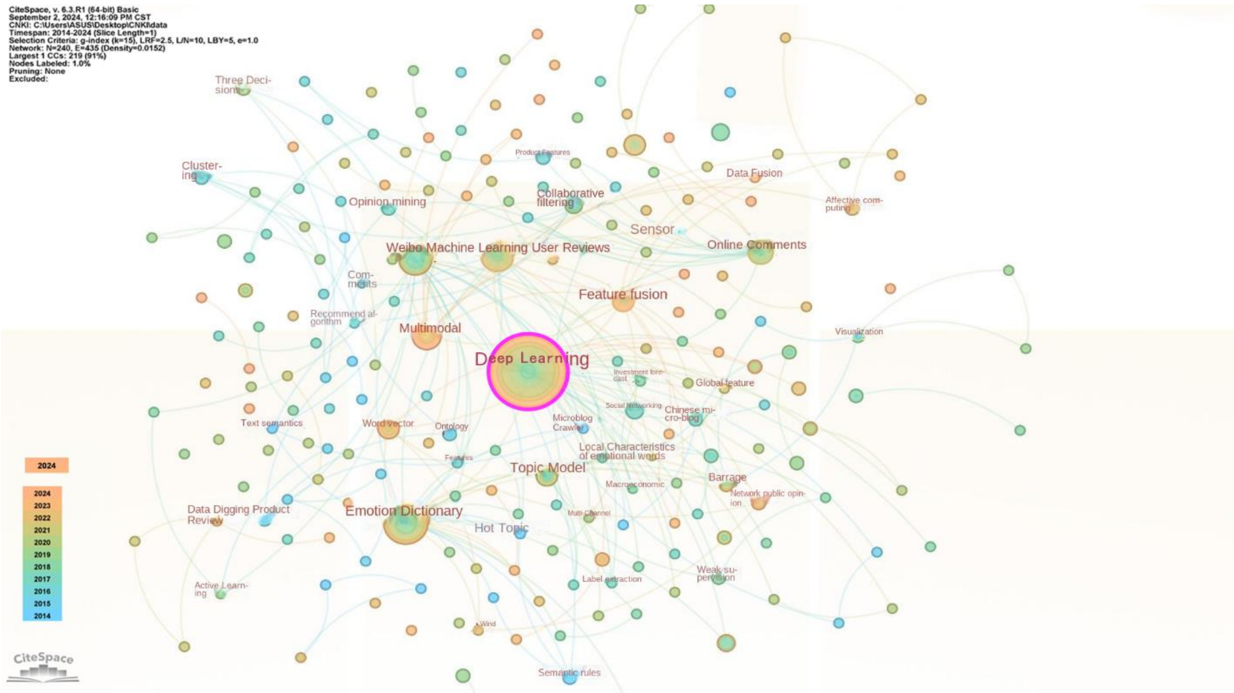

Figure 1. Co-occurrence Map of Keywords in the Field of Sentiment Analysis in Chinese Literature

2.2. International Literature Review

This paper selects the Web of Science Core Collection as the database and uses the advanced search function to construct the search query TS=((sentiment analysis OR emotion analysis OR opinion mining OR sentiment classification) AND (dictionary OR mechanism OR learning OR model OR algorithm)). A preliminary analysis and filtering of the search results in the Web of Science Core Collection were conducted, and 300 articles from the ten-year period between 2014 and 2024 were selected for analysis. The data was analyzed using CiteSpace, and the key thematic keywords in this research field were summarized, as shown in Figure 2.

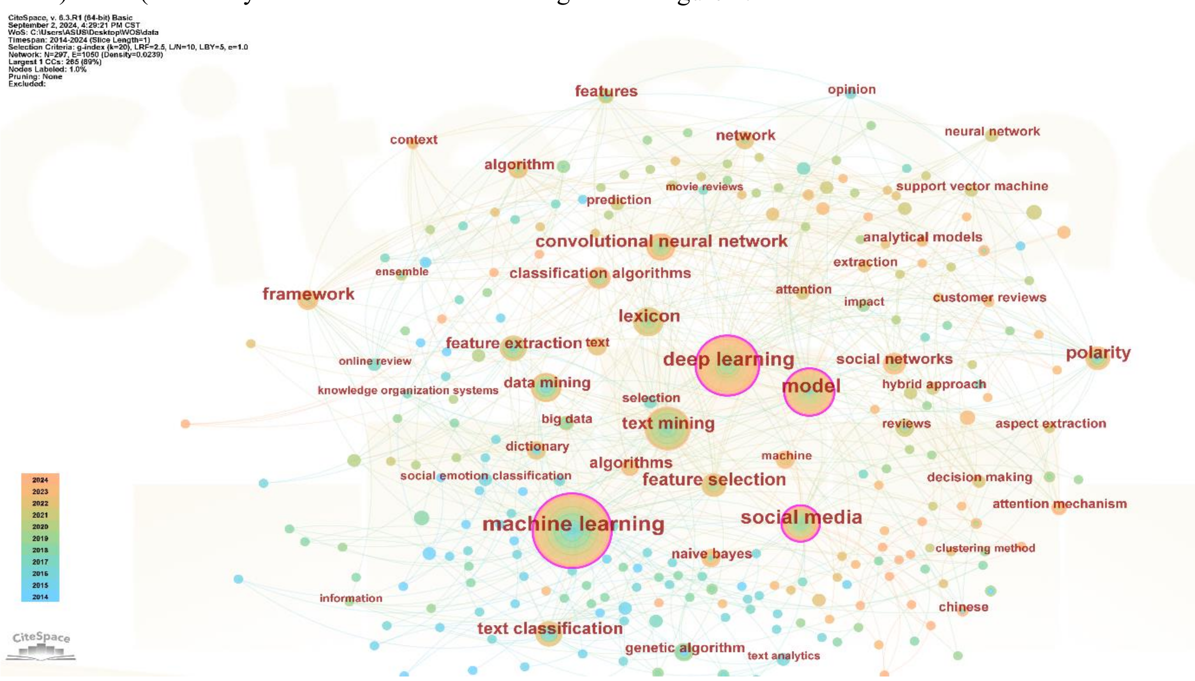

Figure 2. Co-occurrence Map of Keywords in the Field of Sentiment Analysis in International Literature

As shown in Figure 2, the primary application areas of sentiment analysis include social media, online public opinion, and text analysis. The key methods utilized include sentiment lexicon approaches, machine learning techniques, and deep learning strategies. However, internationally, there is more research focused on machine learning methods, while also considering deep learning, along with the introduction of other cutting-edge technologies such as attention mechanisms.

3. Text Sentiment Analysis Methods

3.1. Sentiment Lexicon-Based Methods

The sentiment lexicon-based text sentiment analysis method primarily relies on matching the sentiment polarity of sentiment phrases contained in a sentiment lexicon with the sentiment phrases present in the text. Based on the matching results and the polarity of the sentiment words, the sentiment score is calculated. There is no need to consider the relationships between words; instead, it requires the establishment of corresponding sentiment lexicons for different domains to enhance classification accuracy.

Cai et al. [2] introduced a three-layer sentiment lexicon designed to address the challenge of sentiment words having multiple meanings. Their experiments demonstrated that using a stacked hybrid model, which combines Support Vector Machine (SVM) and Gradient Boosting Decision Tree (GBDT), achieved superior performance compared to baseline single models. To further improve domain specificity in sentiment word selection and accurately determine sentiment polarity within specific domains, Ahmed et al. [3] developed a weakly supervised neural model. This model learns embedded sentiment clusters from the global representation of sentences in the target domain, combining both manual and automated methods. They also proposed an attention-based LSTM model, which extensive experiments on multilingual datasets showed to be effective in improving polarity detection. On the other hand, Czarnek and Stillwell [4] explored the use of two commonly utilized sentiment analysis lexicons, LIWC and NRC, to examine whether older adults express themselves more positively through language. Their findings revealed that relying on a single lexicon might lead to unreliable conclusions, and they recommended using at least two lexicons to ensure accuracy in sentiment analysis.

Although sentiment analysis methods based on sentiment lexicons are relatively simple and easy to implement, they have some significant drawbacks. First, the vocabulary coverage of sentiment lexicons is limited and static, making it difficult to encompass all possible emotional expressions. Language and emotional expression are dynamic, especially in social media and informal texts, where new words, slang, or variants, such as "yin wei ta shan" or "niu ma", may not be recognized. Additionally, this method lacks contextual understanding and struggles to handle complex semantic relationships. There is also the issue of polysemy, where the same word may convey completely opposite sentiments in different contexts, and sentiment lexicons have difficulty correctly distinguishing these cases. To improve the accuracy of sentiment analysis, it is often necessary to combine machine learning or deep learning methods to enhance the analysis results.

3.2. Machine Learning-Based Methods

Machine learning-based text sentiment analysis primarily utilizes three methods: Support Vector Machines (SVM), Naive Bayes, and Maximum Entropy. In their study, Ge et al. [5] applied Naive Bayes and SVM for sentiment analysis and observed that SVM outperformed Naive Bayes in classifying hotel review data. Dhamayanthi N and Lavanya B [6] introduced a novel framework that incorporates machine learning dimensionality reduction (DRML) alongside decision trees, K-nearest neighbors, Bernoulli Naive Bayes, majority voting ensemble, and a custom tokenizer. This framework achieved an average performance of 98.38%, with specific evaluation metrics showing increases of 21.84% in accuracy, 20.4% in precision, 21.84% in recall, and 22.11% in F1-score. Addressing the complexities of Arabic's inflectional and derivational morphology, Ohud Alsemaree et al. [7] analyzed 10,646 Twitter reviews on coffee products using TF-IDF and MRMR, applying four machine learning algorithms: K-nearest neighbors, SVM, decision trees, and random forests. Their ensemble learning analysis demonstrated that hard voting achieved an accuracy exceeding 95.95%, while soft voting reached 94.51%.

Summarizing the above literature, it is evident that machine learning-based sentiment analysis offers advantages over sentiment lexicon-based methods, as it eliminates the need for building and maintaining a sentiment lexicon and allows for learning from full contextual information. However, the effectiveness of machine learning in sentiment analysis heavily depends on the extraction of sentiment features and the selection of appropriate classifiers. This approach requires substantial prior knowledge to properly configure the relevant parameters, which significantly influences the accuracy of the outcomes. Additionally, inappropriate combinations or choices of features and classifiers can adversely affect the final classification accuracy.

3.3. Deep Learning-Based Methods

Deep learning-based sentiment analysis of text is a key area of research within Natural Language Processing (NLP). In contrast to traditional machine learning approaches, deep learning leverages artificial neural networks in multi-layer network architectures, enabling automatic feature extraction and reducing the dependence on manual feature engineering. This approach excels in managing large-scale datasets and complex pattern recognition tasks. The primary deep learning models employed for sentiment analysis include fundamental neural networks like Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory Networks (LSTMs), Bidirectional Long Short-Term Memory Networks (BiLSTMs), and Gated Recurrent Units (GRUs). Researchers often combine these neural networks to further enhance performance and achieve superior results.

Luo et al.[8] proposed a multi-layer network H-RNN-CNN, which introduces a sentence layer into the model and uses a two-layer RNN to model word sequences and sentence sequences, while CNN is used to identify cross-sentence information. This model achieved good results on several Chinese datasets and outperformed all baseline systems in English dataset MR experiments. Guo Xianda et al.[9] proposed an online review sentiment analysis method based on CNN-BiLSTM, which combines the advantages of both neural networks—capturing temporal relationships and local spatial features. Experiments on multi-domain datasets revealed that this method has good domain scalability and can effectively determine consumer sentiment in online reviews, achieving an F1 score of 94.67%.

In recent years, the attention mechanism has been introduced and is widely used in text sentiment analysis due to its ability to effectively capture long-range dependencies, enhance the model's contextual awareness, and improve model interpretability. It is often combined with other models. Hu et al.[10] introduced a multi-layer self-attention mechanism, where BGRU is first used to obtain sequential feature information from the text, followed by initial feature selection using the self-attention mechanism. The processed feature information is then fed into a Convolutional Neural Network (CNN) with different convolutional kernels, and the self-attention mechanism is used again to dynamically adjust the weights of the obtained local features, focusing on the extraction of key features. Finally, Softmax is used to determine the sentiment polarity of the text. The model achieved sentiment classification accuracies of 92.94% and 92.75% on two Chinese corpus datasets, respectively, showing a significant performance improvement compared to mainstream methods.

Spyridon et al. [11] conducted a study on attention-based models built upon Recurrent Neural Networks (RNNs), assessing their effectiveness across various sentiment analysis scenarios. Their research included an exploration of self-attention, global attention, and hierarchical attention techniques, evaluated under different deep neural architectures, training strategies, and hyperparameter settings. For sentiment analysis tasks, the baseline models were compared across experiments both with and without the use of attention mechanisms. The findings revealed that the deep neural networks incorporating attention mechanisms demonstrated excellent performance, highlighting their robustness and scalability in interpreting user opinions and sentiments. Notably, the LSTM model, enhanced with self-attention and word2vec embeddings, delivered the highest accuracy, showing a performance increase of up to 3.5%.

Pre-trained models, which include BERT, GPT, RoBERTa, XLNet, and T5, are used in NLP due to their ability to learn general language representations from large-scale text data. These models are often fine-tuned for specific tasks and combined with other models to create multi-strategy approaches. Abayomi Bello et al. [12] demonstrated that using BERT in combination with Word2vec and deep learning classifiers like CNN, RNN, or BiLSTM achieved state-of-the-art performance in text classification, with accuracy reaching 93% and an F-measure of 95%.Md Saef Ullah Miah et al. [13] proposed an ensemble model using transformers and a large language model for sentiment analysis across multiple foreign languages by first translating them into English using LibreTranslate and Google Translate. They utilized models such as Twitter-Roberta-Base-Sentiment-Latest, bert-base-multilingual-uncased-sentiment, and OpenAI GPT-3 for sentiment analysis. Their results indicated an accuracy of over 86%, demonstrating that translating texts into English enables effective sentiment analysis, and the ensemble model outperformed individual pre-trained models and large language models

4. Conclusion

This paper analyzes and summarizes the recent development trends and major technical methods in sentiment analysis based on the CNKI and Web of Science databases. Using CiteSpace, a large number of relevant domestic and international sentiment analysis publications from the past decade were analyzed, and thematic keywords and timeline maps were generated to identify the main hotspots and technologies over different periods. On this basis, the paper provides a brief introduction and analysis of three mainstream text sentiment analysis techniques, highlighting the strengths and weaknesses of each method. Additionally, it primarily focuses on introducing three cutting-edge sub-methods of deep learning.

When selecting a sentiment analysis method, it is important to make decisions based on specific needs and application scenarios. If the data volume is small, computational resources are limited, and the sentiment analysis task is relatively simple, sentiment lexicon-based methods are the best choice, as they are easy to implement and do not require a large amount of labeled data. For medium-sized data and tasks of moderate complexity, such as product review analysis or general social media sentiment classification, machine learning-based methods can provide a good balance, improving accuracy without demanding excessive resources.However, if faced with large-scale data, highly complex sentiment analysis tasks, and strict accuracy requirements, deep learning-based methods should be chosen. Although they require more computational resources and data, they can handle complex text structures and emotional expressions, providing the most accurate analysis results. Among specific deep learning methods, artificial neural networks (such as LSTM, RNN) are suitable for processing shorter texts or tasks with strong sequential dependencies, especially when contextual relationships are crucial for sentiment judgment, though they may be less effective when dealing with longer texts.If the task involves long texts or requires capturing long-range word dependencies in the text, using self-attention mechanisms (such as Transformers) would be more appropriate, as they can process data in parallel and accurately capture complex semantic relationships. For tasks requiring highly precise sentiment analysis, particularly when the task is complex and data is scarce, pre-trained models (such as BERT, GPT) are the best choice. These models, trained on extensive datasets, possess rich linguistic features and can achieve extraordinary results in specific sentiment assessment activities after fine-tuning.

References

[1] L. Zhang, S. Wang, and B. Liu, “Deep learning for sentiment analysis: A survey,” WIREs Data Min & Knowl, vol. 8, no. 4, p. e1253, Jul. 2018, doi: 10.1002/widm.1253.

[2] Y. Cai et al., “A hybrid model for opinion mining based on domain sentiment dictionary,” Int. J. Mach. Learn. & Cyber., vol. 10, no. 8, pp. 2131–2142, Aug. 2019, doi: 10.1007/s13042-017-0757-6.

[3] M. Ahmed, Q. Chen, and Z. Li, “Constructing domain-dependent sentiment dictionary for sentiment analysis,” Neural Comput & Applic, vol. 32, no. 18, pp. 14719–14732, Sep. 2020, doi: 10.1007/s00521-020-04824-8.

[4] G. Czarnek and D. Stillwell, “Two is better than one: Using a single emotion lexicon can lead to unreliable conclusions,” PLoS ONE, vol. 17, no. 10, p. e0275910, Oct. 2022, doi: 10.1371/journal.pone.0275910.

[5] GE Nilin, FAN Jiajia. Sentiment Analysis of Reviews Based on Naive Bayes and Support Vector Machine[J]. Computer & Digital Engineering,2020,48(7):1700-1704. DOI: 10.3969 / j.issn. 1672-9722.2020.07.029.

[6] D. N and L. B, “A Novel Framework for Sentiment Analysis: Dimensionality Reduction for Machine Learning (DRML),” ijacsa, vol. 15, no. 6, 2024, doi: 10.14569 / IJACSA. 2024. 0150678.

[7] O. Alsemaree, A. S. Alam, S. S. Gill, and S. Uhlig, “Sentiment analysis of Arabic social media texts: A machine learning approach to deciphering customer perceptions,” Heliyon, vol. 10, no. 9, p. e27863, May 2024, doi: 10.1016 / j.heliyon. 2024. e27863.

[8] L. U. O. Fan and W. Houfeng, “Chinese Text Sentiment Classification by H-RNN-CNN,” Acta Scientiarum Naturalium Universitatis Pekinensis, vol. 54, no. 3, p. 459, May 2018, doi: 10.13209 /j.0479-8023.2017.168.

[9] GUO Xianda, ZHAO Narisa, CUI Shaoze. Consumer reviews sentiment analysis based on CNN-BiLSTM. Systems Engineering - Theory & Practice, 2020, 40(3): 653-663 https://doi.org/10.12011/1000-6788-2018-1890-11

[10] [HU Yan-li, TONG Tan-qian, ZHANG Xiao-yu, PENG Juan. Self-attention-based BGRU and CNN for Sentiment

Analysis[J]. Computer Science, 2022, 49(1): 252-258. https://doi.org/10.11896/jsjkx.210600063

[11] S. Kardakis, I. Perikos, F. Grivokostopoulou, and I. Hatzilygeroudis, “Examining Attention Mechanisms in Deep Learning Models for Sentiment Analysis,” Applied Sciences, vol. 11, no. 9, Art. no. 9, Jan. 2021, doi: 10.3390/app11093883.

[12] A. Bello, S.-C. Ng, and M.-F. Leung, “A BERT Framework to Sentiment Analysis of Tweets,” Sensors, vol. 23, no. 1, Art. no. 1, Jan. 2023, doi: 10.3390/s23010506.

[13] M. S. U. Miah, M. M. Kabir, T. B. Sarwar, M. Safran, S. Alfarhood, and M. F. Mridha, “A multimodal approach to cross-lingual sentiment analysis with ensemble of transformer and LLM,” Sci Rep, vol. 14, no. 1, p. 9603, Apr. 2024, doi: 10.1038/s41598-024-60210-7.