Mathematical Modeling and Algorithm Application
ISSN: 3006-0842 | Vol. 2, No. 2, 2024

Analysis of Online Public Opinion Texts Based on Topic Mining and Sentiment Analysis

Bo He, Lei Wang*, Ruoyu Zhao, Yongfen Yang

School of Computer Science and Engineering, Chongqing University of Technology, Chongqing, 400054, China
* **Corresponding author**: Lei Wang (Email: wanglei6731@163.com)

Abstract: Text-based public opinion analysis is an important branch of natural language processing. It aims to collect, analyze, and interpret the opinions, emotions, and attitudes of the general public to understand and predict their reactions to specific events, policies, or companies. In recent years, this field has become a research hotspot. This paper reviews methods of public opinion analysis based on topic mining and sentiment analysis, exploring their concepts and characteristics and analyzing recent research achievements. By comparing the advantages and disadvantages of different methods, this paper summarizes the strengths and limitations of these approaches. Based on a review of the current state of research both domestically and internationally, this paper provides an in-depth analysis of public opinion analysis methods and proposes future directions and trends for development.

Keywords: Text-based public opinion analysis; sentiment analysis; topic mining.

1. Introduction

In today's rapidly developing information and globalized society, the widespread adoption of the internet has led to the rise of numerous social media platforms. According to statistics from Statista, as of 2023, the number of internet users worldwide has exceeded 5.3 billion, with social media users reaching 4.4 billion, who spend an average of over two hours daily on social media. This vast user base generates a massive amount of information and data on the internet every day, providing abundant material for public opinion analysis. For instance, over 500 million tweets are posted on Twitter daily, and Facebook has more than 2 billion daily active users. These figures highlight the significant role of the internet and social media in information dissemination and underscore the importance of natural language processing (NLP) in the current social development.

NLP is a discipline that studies the theories and methods for effective communication between humans and computers using natural language. As an important branch of NLP, public opinion analysis focuses on analyzing and understanding the statements and viewpoints people express on social media. It has become one of the hot research directions in the field of NLP. Through public opinion analysis, governments, businesses, and individuals can better understand the public's attitudes and emotions toward specific topics or events, enabling them to respond or adjust strategies promptly. Additionally, public opinion analysis can be used to predict market trends, monitor public opinion, identify potential risks, and has broad applications in politics, commerce, society, and other fields.

2. Public opinion analysis

Public opinion analysis is the process of systematically collecting, analyzing, and interpreting social opinions and emotions to gain insights into the public's attitudes and emotional tendencies toward specific topics or events. In the study of public opinion on social media, domestic and international scholars have different focuses. International researchers tend to emphasize the monitoring and technical aspects of public opinion on social media. For example, Keela et al. [1] classified and analyzed the emotional trends of the public by extracting features from social media comment texts, concluding that governments should intervene in the development of social media opinion in a timely manner and take effective measures. Ma et al. [2] explored the origin of public opinion by mining its propagation patterns and proposed strategies for responding to public opinion. Domestic researchers primarily focus on the evolution, monitoring, and guidance of public opinion over time. In the study of opinion evolution, there is a greater emphasis on the evolution of topics. For instance, Zhu Xiaoxia et al. [3] proposed a method for public opinion analysis that integrates dynamic topics and sentiments, dynamically tracking the evolution of topics at different stages. Shao Qi et al. [4] explored network public opinion topics of public events from different semantic perspectives, using PageRank to identify core nodes and uncover deep topic evolution information.

Text mining and NLP are both crucial technologies in public opinion analysis. Numerous scholars have applied text-mining techniques in public opinion research and have made significant progress [5]. Text mining involves multiple processes to extract valuable information from text data. Natural language processing is an essential technology for computer processing of natural language, aimed at analyzing text data sets to accomplish various text-related tasks.

With the widespread adoption of social media in public communities, public opinion analysis has undergone significant changes, no longer relying solely on traditional surveys and interviews. Online opinion analysis based on social media data has garnered significant interest from scholars and practitioners, as social media platforms are considered valuable sources of information or online insights [6]. Public opinion based on social media information provides an excellent opportunity to explore the public's mindset on specific topics. This makes opinion analysis an ideal tool for extending product analysis, managing organizational reputation, public relations, predicting sports event outcomes, and more. Generally, sentiment analysis and topic analysis are two crucial components of online public opinion analysis. Currently, online public opinion analysis can be divided into topic-based analysis and sentiment-based analysis, depending on the methods used. Typically, these two approaches are combined: data preprocessing of opinion texts is conducted, followed by topic mining and sentiment analysis, and finally, a comprehensive analysis is performed by integrating the results of opinion hotspots and sentiment classification. The specific methodology is illustrated in Figure 1.

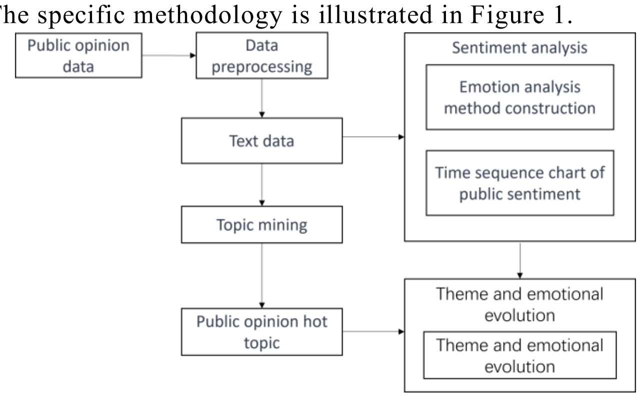

Figure 1 flow chart of public opinion analysis by combining topic mining and sentiment analysis

3. Network public opinion text analysis based on topic mining

In public opinion analysis, topic mining plays a crucial role. By combining these two techniques, it is possible to comprehensively understand and analyze large amounts of text data, thereby providing strong support for decision-making. Topic mining, also known as topic discovery or topic extraction, is a technique used to automatically label and extract representative words, phrases, or sentences. It can quickly and effectively identify interesting and valuable information from massive data, especially demonstrating unique advantages in processing large-scale texts. Several common methods of topic mining include frequency-based methods, semantic analysis-based methods, topic probability model-based methods, and machine learning-based methods.

3.1. Frequency-based methods

Luhn from IBM proposed the automatic indexing method based on word frequency statistics, which initiated the research on topic extraction. TF-IDF (Term Frequency-Inverse Document Frequency) is currently the most commonly used technique for extracting topic words, but its application is limited to relatively stable topic expressions. Therefore, some scholars propose to discriminate and filter topic words by examining the statistical features of the words themselves, such as IDF (Inverse Document Frequency), Part-of-Speech (POS), and word position in the document. Based on this, Han Kesong and Wang Yongcheng designed a weighted system for extracting topic words based on the actual language expression situation. In addition, some scholars extract topic words by examining the association information between words, such as mutual information, word span, and co-occurrence relationship. Considering the characteristics of news text, literature selects relevant features (including position, title sentence relevance, etc.) and establishes a topic sentence extraction model through feature weighting. Although the frequency-based method is simple and straightforward and does not require training models, its weighting system may change with different text characteristics, resulting in weak adaptability to different domains.

3.2. Semantic Analysis Methods

Methods based on semantic analysis can deeply analyze the semantics, grammar, and topic distribution in text and mainly include the following methods: ontology, lexical chains, semantic dictionaries, and shallow semantic analysis. Zhu Hengmin and Ma Jing introduced domain ontology to represent text topics, effectively reducing the dimensionality of text features. Lexical chains were first proposed by Halliday and Hasan, who believed that there is a certain semantic coherence between words under the same topic. By constructing lexical chains, words scattered in the original text are transformed into sets of semantically related vocabulary, which can effectively reflect topic information. Methods based on semantic dictionaries utilize pre-constructed dictionaries to understand the semantic relationships in the text, thereby improving the accuracy of topic extraction.The shallow semantic analysis identifies words and phrases related to topics by analyzing the shallow semantic structure of sentences. Methods for topic extraction based on semantic analysis are usually used as auxiliary methods, combined with other methods to effectively improve extraction accuracy.

3.3. Methods based on topic probability models

The topic probability model is an unsupervised method for topic extraction that effectively mines the latent semantic information in corpora. The two most common topic probability models are Probabilistic Latent Semantic Analysis (PLSA) and Latent Dirichlet Allocation (LDA). PLSA is a topic probability model built on maximum likelihood and generative models. It assumes that a document consists of multiple topics, and topics are probability distributions of all vocabulary words. The probability of each word appearing in a document is calculated as follows:

$$P\left(\frac{words}{documents}\right) = \sum P\left(\frac{words}{theme}\right) * P\left(\frac{theme}{documents}\right) \tag{1}$$

PLSA effectively addresses the problem of polysemy in the traditional Latent Semantic Analysis (LSA) model. However, due to its large number of parameters, it may suffer from overfitting. To address this issue, Blei introduced the Dirichlet prior distribution. The Latent Dirichlet Allocation (LDA) model is a typical three-layer Bayesian model, as illustrated in Figure 2. Its basic idea is to represent the document-topic and topic-word relationships as multinomial distributions with Dirichlet prior probabilities. LDA addresses the overfitting problem of PLSA by introducing hyperparameters.

Although topic models have made significant progress in topic mining, as a generative model, their scalability is limited, and their semantic expression is incoherent. Combining semantic-based methods with topic models or incorporating domain knowledge into topic models is one of the directions for future improvement.

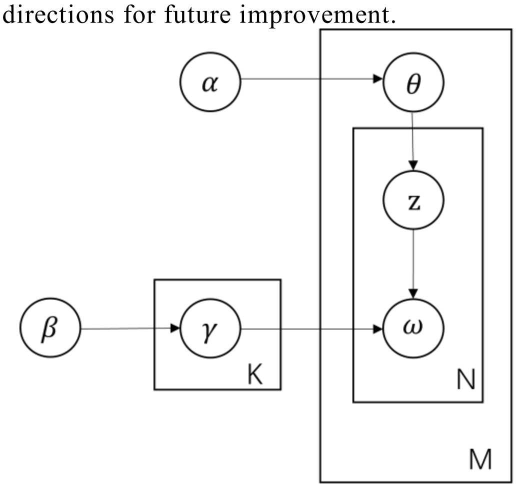

Figure 2 LDA topic model diagram

3.4. Methods based on machine learning

Methods based on machine learning transform the problem of automatic topic extraction into a topic classification problem. Turney and Witten et al. designed the GenEx and Kea systems using C4.5 decision trees and naive Bayes methods for extracting key phrases and topic words, achieving accuracies of 23.9% and 33.27%, respectively. However, relying solely on machine learning algorithms may not effectively understand the content of the text. Therefore, Quan proposed a statistical machine learning approach, which combines five classification algorithms (naive Bayes, support vector machine, logistic regression, random forest, ensemble methods) with five statistical features (word frequency, inverse document frequency, betweenness centrality, TextRank, co-occurrence) to compare and analyze the performance of each combination in topic extraction. Due to the strong flexibility of neural network models in handling different datasets, some scholars have introduced them into topic extraction with good results.

3.5. Comparison of Topic Mining Methods

By summarizing several common topic extraction techniques in current research, this paper outlines their respective advantages and disadvantages as well as future development trends, as shown in Table 1.

Table 1 The advantages and disadvantages of topic mining analysis method and its development direction

Methods	Advantages and disadvantages	Development trends
Frequency-Based methods	Simple and straightforward, no need for model training;Does not consider relationships between words and semantic features	Mature development, widely applied, often combined with other methods
Semantic analysis methods	Considers word relationships, enhancing semantic coherence; Narrow applicability	Combining with other methods to improve topic extraction accuracy
Methods based on topic probability models	Strong portability; Poor semantic coherence, insufficient feature expression ability, and complex generative model	Mainstream method, widely applicable across multiple domains
Machine learning methods	Improving extraction accuracy by integrating multiple features; Dependent on data quality and feature selection effectiveness	Mature development, widely applied; Deep learning frameworks are one of the future directions

4. Network public opinion text analysis based on sentiment analysis

Predicting the evolution of sentiment in public opinion events is also one of the research directions in sentiment analysis. Researchers typically employ analysis algorithms including machine learning to analyze sentiment in public opinion data. By using appropriate algorithms to classify or cluster sentiment tendencies, it becomes clearer to identify the sentiment orientation of public opinion. Sentiment analysis, also known as opinion mining and polarity analysis, typically refers to the process of extracting, analyzing, summarizing, and inferring subjective texts with emotional color. Several common methods of sentiment analysis include lexicon-based methods, machine learning-based methods, and deep learning-based methods. Currently, sentiment analysis finds wide applications in areas such as public opinion monitoring, product reviews, and event detection.

4.1. Sentiment Lexicon based approach

The model based on the sentiment lexicon utilizes a large number of vocabulary to construct and train a sentiment dictionary. Then, it categorizes the words based on their sentiment scores, and the final classification effectiveness depends on the completeness of the sentiment lexicon. The flowchart of this method is shown in Figure 3.

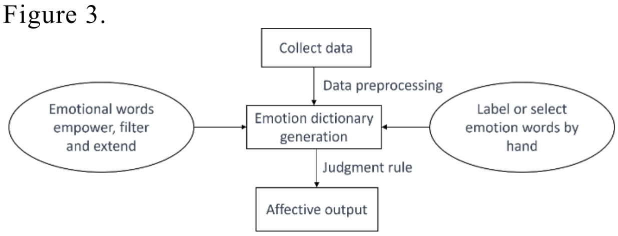

Figure 3 flow chart of sentiment analysis method based on emotion dictionary

The earliest English sentiment lexicon to appear overseas is SentiWordNet. Commonly used sentiment lexicons include General Inquirer, Opinion Lexicon, and

MPQA (Multi-Perspective Question Answering). Unlike English sentiment lexicons, Chinese sentiment lexicons mainly consist of NTUSD, HowNet, and the sentiment lexicon ontology library. These three types of sentiment lexicons contain varying numbers of positive and negative words, and many sentiment analysis researchers have extensively studied and used them. However, lexicon-based methods face challenges in terms of complex and time-consuming construction processes and ineffective performance when dealing with newly emerged internet slang. In contrast, machine learning-based methods utilize known data to predict the sentiment of unknown data, offering higher classification accuracy and generalization performance.

4.2. Methods based on machine learning

Traditional machine learning methods not only effectively address the limitations of sentiment lexicons in application but also offer relatively higher classification accuracy and stronger generalization performance. Shallow machine learning methods involve selecting sentiment words as features, annotating, and matrixing text to form a text feature set. Text feature extraction is generally based on the Term Frequency-Inverse Document Frequency (TF-IDF). Subsequently, classification models constructed using logistic regression, naive Bayes, and support vector machine (SVM) are used to predict sentiment tendencies in the test dataset. The classification effectiveness depends on the selection of training texts and accurate sentiment labeling. The data processing flow based on machine learning is illustrated in Figure 4.

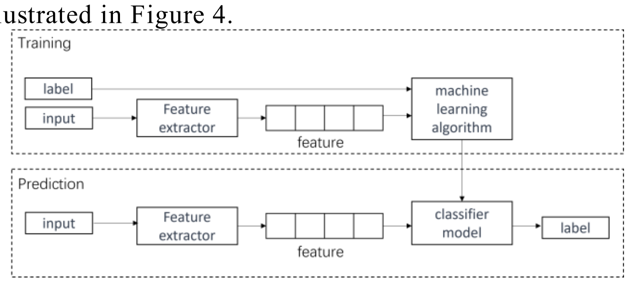

Figure 4 Sentiment analysis method diagram based on machine learning

Machine learning methods, such as Support Vector Machine (SVM) and Naive Bayes (NB), have been widely applied in sentiment classification research. Wang et al. proposed a novel aspect-based sentiment analysis method that achieves precise machine annotation without manual labeling. Text sentiment analysis overly relies on the statistics of sentiment words, and the issue of inadequate consideration of vocabulary and latent semantic information in SVM sentiment analysis has not been fully addressed. Han et al. introduced an improved Support Vector Machine based on the Fisher kernel function (FK-SVM). However, machine learning heavily relies on feature representation, and due to the complexity of feature engineering, it is challenging to obtain satisfactory classification results. Faced with massive data resources, feature annotation requires a significant amount of manpower, making it time-consuming and labor-intensive.

4.3. Methods based on deep learning

Methods based on deep learning primarily utilize artificial neural networks to avoid manually constructing a large number of text features, thus saving manpower and resources, effectively overcoming the limitations of shallow machine learning. Popular deep-learning networks include CNN, RNN, LSTM, GRU, BERT, and attention mechanisms.

Basiri et al. [36] proposed sentiment analysis based on an attention-based bidirectional CNN-RNN deep model. This study utilized both past and future contexts and considered five reviews and three Twitter datasets. Jin et al. [37] introduced sentiment analysis based on heterogeneous graphs, utilizing network embeddings based on variational autoencoders to learn joint representations of user social relationships. This was encouraged by preserving both structural proximity and attribute proximity, and the model outperformed traditional text-based sentiment analysis methods. Pota et al. [38] recommended a pipeline for Twitter sentiment analysis based on Bidirectional Encoder Representations from Transformers (BERT). This study is interesting as it aimed to transform slang into plain text, using BERT classification on tweets but pre-trained on plain text. The model is applicable to multiple languages. Nemes and Kiss [39] conducted sentiment analysis on social media based on COVID-19 (comments, tags, posts, tweets). Although the COVID-19 pandemic has had a global impact [40-41], this study considered the use of Recurrent Neural Networks (RNNs) for analysis. The research concluded that there were more positive tweets on social media.

In recent years, Explainable Artificial Intelligence (XAI) methods have been introduced in various fields and domains to further validate and interpret deep learning models. Abdelwahab et al. (2022) [42] proposed an attention-based Long Short-Term Memory (LSTM) approach for Explainable Arabic Sentiment Analysis (ASA), demonstrating how LSTM leads to emotional polarity prediction in ASA on Twitter users' medical opinions about LASIK surgery in specific domains of Arabic text. Mohana & Rajathi (2022) [43] introduced a sentiment analysis method for electronic shopping product text based on Chaos Coyote Optimization Deep Belief Networks to enhance sentiment analysis accuracy. The main goal of this method is to classify text sentiment based on polarity (positive and negative) and improve accuracy. Liu et al. (2023) [44] utilized text sentiment analysis to study public sentiment changes during the COVID-19 pandemic, using the SNOWNLP module to compute sentiment scores of Weibo comment text data and visualize sentiment score distributions. On the other hand, sentiment classification based on the Support Vector Machine algorithm was performed, comparing models trained with three different kernel functions, with results indicating that the polynomial kernel function performs the best, demonstrating good classification performance of the sentiment classifier.

4.4. Comparison of Sentiment Analysis Methods

By comparing and analyzing methods based on sentiment lexicons, machine learning, and deep learning, this study aims to summarize their respective advantages and disadvantages, as shown in Table 2.

Table 2 The advantages and disadvantages of text sentiment analysis methods

Methods	Advantages	Disadvantages
Sentiment Analysis Method Based on Sentiment Lexicon	It effectively reflects the structural features of the text, is easy to understand, and shows significant sentiment classification effects when there are sufficient sentiment words.	It does not overcome the limitations of the sentiment lexicon and requires continuous expansion of the lexicon, resulting in a lower accuracy of text sentiment judgment.
Sentiment Analysis Methods Based on Machine Learning	It can classify the sentiment of text based on the selection of sentiment features and the combination of sentiment classifiers.	These methods are unable to capture contextual information from the surrounding text.
Sentiment Analysis Methods Based on Deep Learning	It can actively learn textual features, preserving the sequential information of words in the text. By utilizing deep neural network models to learn key information from the data, it reflects the data's characteristics, thus enhancing learning performance.	This method requires substantial data support and is not suitable for small-scale datasets, often resulting in longer training times for the algorithm.

5. Conclusion

Network public opinion analysis is one of the hottest research areas in natural language processing and has been widely applied in various smart city applications. Past studies have emphasized the importance of sentiment analysis in public opinion analysis and have employed various traditional and non-traditional methods, including artificial intelligence technologies. In the future, we hope to delve deeper into this issue by leveraging hybrid machine-learning techniques and other hyperparameter optimization methods. Furthermore, we can attempt to introduce more modal content to improve the model's performance, such as news comments, background information on news releases, and news video content. Due to limitations of existing public opinion news datasets, the adaptability of our model to other datasets remains to be validated, thus requiring training on larger-scale and diverse real-world datasets to enhance its generalization ability. Additionally, it is unreasonable to simply categorize the task of sentiment tendency detection in public opinion news as a binary classification problem. Although a small portion of studies have conducted fine-grained sentiment analysis, data related to public opinion is relatively scarce. Future research could further analyze news in a more detailed manner, attempting to transform it into multi-class or even regression tasks, which also poses a major challenge for future work.

Acknowledgements

This research is supported by the chongqing university of action plan for high quality development of postgraduate education of chongqing university of technology under grant No. gzlcx20243211.

References

[1] Keelan J, Pavri V, Balakrishnan R, et al. An analysis of the Human Papilloma Virus vaccine debate on MySpace blogs[J]. Vaccine, 2010, 28(6): 1535-1540.

[2] Ma Y, Shu X, Shen S, et al. Study on network public opinion dissemination and coping strategies in large fire disasters[J]. Procedia engineering, 2014, 71: 616-621.

[3] Zhu, X X., Song, J X, Meng, J F. Network public opinion information analysis based on dynamic topic-sentiment evolution model. [J]. Information Science, 2019, 37(07): 72-78.

[4] Shao Q, Mou D M, Wang P, Jin C Y. Research on Semantic-based Topic Discovery of Network Public Opinion on Sudden Public Health Events[J]. Data Analysis and Knowledge Discovery, 2020, 4(09): 68-80..

[5] Zheng Y, Qin Y Y, Shi Y H,et al. Analysis of Current Research Status of Network Public Opinion[J]. Industrial Control Computer, 2023, 36(07): 100-102.

[6] Wan X, Wang R, Wang M, et al. Online public opinion mining for large cross-regional projects: Case study of the south-to-north water diversion project in China[J]. Journal of Management in Engineering, 2022, 38(1): 05021011.

[7] LUHN H P. A Statistical approach to mechanized encoding and searching of literary information [J]. IBM Journal of Research and Development,1957, 1(4): 309-317.

[8] SALTON G, BUCKLEY C. Term weighting approaches in automatic text retrieval[J]. Inforation Processing & Management, 1987, 24(5): 513-523.

[9] CHEN Y H，LU J L，MENG F T. Finding keywords in blogs: efficient keyword extraction in blog mining via user behaviors[j]. Expert Systems with Applications an International Journal, 2014, 41(2): 663-670.

[10] BOKAETF M H, SAMETI H, LIU Y. Unsupervised approach to extract summary keywords in meeting domain [C] / /Signal Processing Conference. IEEE, 2015: 1406-1410.

[11] AWAJAN A. Keyword extraction from Arabic documents using term equivalence classes [j]. ACM Transactions on Asian and Low-Resource Language Information Processing, 2015, 14(2): 1-18.

[12] Han K, Wang Y C. A Nonlinear Weighting Method for Topic Extraction[J]. Journal of the China Society for Scientific and Technical Information, 2000, 19(6): 650-653.

[13] KOUYING X,SHENG Y,XINYAN Z. Topic sentence extraction method based on weight fuzzy clustering and mutual information[j]. Computer Engineering,2009,35 (20): 184- 186.

[14] Zhang Jin. Information Keyword Extraction Method Based on Improved TF-IDF Algorithm[J]. Journal of Intelligence, 2014, 33(4): 153-155.

[15] Ruan G C, Xia L. Research on Text Topic Keyword Clustering Based on Co-occurrence Analysis[J]. Journal of Library Science, 2018, 37(11): 99-104.

[16] Pu M, Zhou F, Zhou J J, et al. News Key Event Topic Sentence Extraction Based on Weighted TextRank[J]. Computer Engineering, 2017, 43(8): 219-224.

[17] Wan G, Zhang G P, Bai Y, et al. News Topic Sentence Extraction Based on Feature Weighting[J]. Journal of Chinese Information Processing, 2017, 31(5): 125-131.

[18] He C, Zhang Y F. Research on Chinese Text Semantic Feature Extraction Algorithm Based on Fusion of Domain Ontology[J]. Information Studies: Theory & Application, 2013, 36(9): 96-99.

[19] HALLIDAY M A K,HASAN R. Cohesion in English[M]. London: Routledge, 1976: 6-11.

[20] Liu M, Wang X L, Liu Y C. Research on Key Phrase Extraction Method Based on Lexical Chain[J]. Chinese Journal of Computers, 2010, 33(7): 1246-1255.

[21] Zeng C, Zhang D Z. Phrase Topic Extraction Based on Synonym Concept Tree and "HowNet" [J]. Journal of Xiamen University: Natural Science Edition, 2015, 54(2): 263-269.

[22] Zhang F, Le X Q. Research on Method of Identifying Topic Attributes Instances in Innovation Sentences of Domain Scientific Literature[J]. Modern Library and Information Technology, 2015, 31(5): 15-23.

[23] HOFMANN T. Unsupervised learning by probabilistic latent semantic analysis[J]. Machine Learning, 2001, 42 (1-2): 177-196.

[24] BLEI D M, NG A Y, JORDAN M I. Latent Dirichlet allocation[J]. Journal of Machine Learning Research, 2003, 3: 993-1022.

[25] TURNEY P D. Learning to Extract Keyphrases from Text, ERB1057[R]. Washington DC: National Research Council, 1999.

[26] WITTEN I H, PAYNTER G W, FRANK E, et al. KEA: practical automatic keyphrase extraction [C] //Proc of the 4th ACM Conference on Digital Libraries. New York: ACM Press, 1999: 254-256.

[27] ONAN A,KORUKOGLU S,BULUT H. Ensemble of keyword extraction methods and classifiers in text classification [J]. Expert Systems with Applications, 2016, 57(C): 232-247.

[28] Wang T, Yang W Z. A Review of Methods for Text Sentiment Analysis[J]. Computer Engineering and Applications, 2021, 57(12): 11-24.

[29] Baccianella S, Esuli A, Sebastiani F. Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining[C]//Lrec. 2010, 10(2010): 2200-2204.

[30] Ku L W, Chen H H. Mining opinions from the Web: Beyond relevance retrieval[J]. Journal of the American Society for Information Science and Technology, 2007, 58(12): 1838-1850.

[31] Cahyani D E, Patasik I. Performance comparison of tf-idf and word2vec models for emotion text classification[J]. Bulletin of Electrical Engineering and Informatics, 2021, 10(5): 2780-2788.

[32] Xu G, Yu Z, Yao H, et al. Chinese text sentiment analysis based on extended sentiment dictionary[J]. IEEE Access, 2019, 7: 43749-43762.

[33] Ortis A, Farinella G M, Battiato S. Survey on visual sentiment analysis[J]. IET Image Processing, 2020, 14(8): 1440-1456.

[34] Wang Y, Chen Q, Shen J, et al. Aspect-level sentiment analysis based on gradual machine learning[J]. Knowledge-Based Systems, 2021, 212: 106509.

[35] Han K X, Chien W, Chiu C C, et al. Application of support vector machine (SVM) in the sentiment analysis of twitter dataset[J]. Applied Sciences, 2020, 10(3): 1125.

[36] Basiri ME, Nemati S, Abdar M, Cambria E, Acharya UR (2020) ABCDM: an attention-based bidirectional CNN–RNN deep model for sentiment analysis. Future Gener Comput Syst 115: 279-294.

[37] Jin Z, Zhao X, Liu Y (2021) Heterogeneous graph network embedding for sentiment analysis on social media. Cogn Comput 13(1):81-95.

[38] Pota M, Ventura M, Catelli R, Esposito M (2021) An effective BERT-based pipeline for Twitter sentiment analysis: a case study in Italian. Sensors 21(1): 133.

[39] Nemes L, Kiss A (2021) Social media sentiment analysis based on COVID-19. J Inf Telecommun5(1):1-15.

[40] Priyadarshini I, Mohanty P, Kumar R, Son LH, Chau HTM, Nhu VH, Tien Bui D (2020) Analysis of outbreak and global impacts of the COVID-19. In: Healthcare, vol 8(2). Multidisciplinary Digital Publishing Institute, p 148.

[41] Dansana D, Kumar R, Das Adhikari J, Mohapatra M, Sharma R, Priyadarshini I, Le DN (2020) Global forecasting confirmed and fatal cases of covid-19 outbreak using autoregressive integrated moving average model. Front Public Health 8: 580327.

[42] Abdelwahab Youmna,Kholief Mohamed,Sedky Ahmed Ahmed Hesham. Justifying Arabic Text Sentiment Analysis Using Explainable AI (XAI): LASIK Surgeries Case Study[J]. Information, 2022, 13(11).

[43] Mohana R S,Rajathi K,Kousalya K,Yuvaraja T. Text sentiment analysis on E - shopping product reviews using chaotic coyote optimized deep belief network approach[J]. Concurrency and Computation: Practice and Experience,2022,34(19).

[44] Liu H H, Wang A Y, Liu Y T. Text Sentiment Analysis Based on SVM: A Case Study of the COVID-19 Pandemic[J]. Information Technology and Informatization, 2023, (01): 37-40.