Frontiers in Computing and Intelligent Systems
ISSN: 2832-6024 | Vol. 3, No. 1, 2023

Research on sentiment analysis methods for text-oriented data

Yuanfei Deng

School of Computer Science, South China Normal University, Guangzhou 510631, China
dengyf@m.scnu.edu.cn

Abstract: With the rapid development of information technology, the results obtained from sentiment analysis on a large number of speech information on these platforms can be used for comment classification, product analysis and recommendation, consumption forecast and other aspects of the network platform With the rapid development of information technology, the results obtained from sentiment analysis on a large number of speech information on these platforms can be used for comment classification, product analysis and recommendation, consumption forecast and other aspects of the network platform Sentiment analysis is a practical technique which has become one of the most active research fields in natural language processing. The traditional text sentiment analysis method consumes a lot of human resources, but the coverage of artificial extracted features is the traditional text sentiment analysis method consumes a lot of human resources, but the coverage of artificial extracted features is limited and the artificial irrational behavior will affect the correctness of the results, so the traditional method is not universal. With the development of deep learning, text pre-training language model and knowledge graph technology continue to develop. Aiming at the research of sentiment analysis methods for text data, we summarize the research background and domestic and foreign research status of sentiment analysis methods for text data, and explore the hot research content, key problems, commonly used experimental methods and technical lines of sentiment analysis methods for text text data.

Keywords: Sentiment analysis; Aspect level; Knowledge graph; Pre-training model.

1. Introduction

With the rapid development and popularity of the Internet and information technology, the scale and number of various computer software, websites, mobile phone APPs and small programs are rapidly expanding, and the data generated by them is exploding, and the world has entered the Internet + big data era. 31 August 2015, the State Council issued the "Action Plan for Promoting the Development of Big Data", which mentioned the need to promote the integration of big data resources and promote big data industry innovation and development to help social and economic transformation and upgrading[7] On 29 March 2018, the White Paper on Big Data Standardisation (2018 Edition), edited by the China Institute of Electronic Technology Standardisation, was officially released[8] to promote the continuous development and improvement of big data- related technologies, products, applications and standards, further highlighting the technical support and leading role of big

data standards. In the era of big data, the accumulation of massive amounts of text continues to emerge in various fields. Text (natural language) is the most natural way to encode human knowledge, and is the most common type of information people encounter, the most expressive form of information. From humanities research to government decision-making, from precision medicine to quantitative finance, from customer management to marketing, massive text plays a pivotal role as one of the most important information carriers everywhere.

Sentiment analysis is an active research area in text mining. It identifies, extracts, quantifies and processes the sentiment states and subjectivity of texts in a computational process and systematic way. Among the different applications of SA, it is crucial to gain insight into public opinion on various socio-political topics by analysing tweets and other social media text public material, automatically analysing historical corpora, and studying product reviews for real customer feedback and customer sales forecasts[1] . Due to the downward trend in data storage prices and the rapid growth of information technology, the global community stores huge amounts of data every day. One of the main focuses of natural language processing (NLP) is the design and construction of computational platforms using artificial intelligence (AI) methods[2][3][4][5] . These platforms automate the process of extracting knowledge and previously unknown interesting patterns from structured and unstructured text sources[6] . Sentiment analysis is a method for extracting the contextual polarity (positive, negative or neutral) of texts.

Text sentiment analysis focuses on how to enable computers to automatically identify, classify, annotate or extract the subjective content of human sentiment, emotions, opinions, evaluations, etc. expressed in natural language texts about a given subject or topic[9][10] The sentiment analysis is the evaluation of the public's opinions, emotionally recognized attitudes towards products, services, organizations and other objects[11] The analysis of the sentiment information in the review text allows the extraction of the user's emotional attitude expressed in a piece of review text[12] In recent years, as social networks and e-commerce have become more and more connected to each other, it is possible to extract the emotional attitude of users in a review text. In recent years, with the rapid development of social networks and e-commerce platforms, the Internet has generated a huge amount of text data, which requires the use of computers to process the massive amount of information that cannot be processed manually. How to compute sentiment on the massive amount of text data from e-commerce platforms and social networks and uncover the hidden value behind it will have important commercial value and research significance. Sentiment analysis has the need to organise and analyse text at both the whole document and semantic detail levels.

Sentiment analysis is a practical technique that has become one of the most active areas of research in natural language processing (NLP). Much progress has been made with the help of existing research results in NLP[13] Emotion Analysis Sentiment analysis allows companies, researchers, governments, politicians and organizations to understand people's emotions, which play an important role in the decision-making process. To enable computers to mimic humans, the focus and difficulty of research is the understanding and analysis of human emotions, which is an important task in the development of artificial intelligence. The results of emotion analysis research can be widely used for needs such as online opinion monitoring, hot topic tracking and product review classification[14] . In e-commerce practice, sentiment analysis of user reviews can help vendors improve their products and after-sales services[15][16] This is essential for product improvement and decision making.

The main contributions of this paper are as follows.

(1) An overview summarizing the research background and the current status of domestic and international research on sentiment analysis methods for text-oriented data.

(2) To point out the hot research contents, key problem studies, common experimental methods and technical lines of research on sentiment analysis methods for text-oriented data.

2. Current status of domestic and international research

In recent years, research related to affective computing has gained increasing attention in academia and industry, and has become a popular research direction in areas such as data mining and deep learning[17] . Emotion computing is one of the research areas of natural language processing, which has been discussed in top international conferences such as ACL, EMNLP, NACCL, COLING, AAAI, IJCAI, ICML, NIPS[18] . Due to the rapid development of computer technology, the fields of artificial intelligence and natural language processing place more emphasis on publishing conference papers and being able to communicate the latest advances through conferences in a timely manner. However, there are also exclusive academic journals in top international journals such as Computational Linguistics, Transactions of the Association for Computational Linguistics, ACM Transactions on Speech and Language Processing, etc. The top international journals also have their own academic journals. In China, there are numerous research units focusing on cutting-edge research in multi- emotional computing, including the Social Computing and Information Retrieval Laboratory at HUST, the Natural Language Processing and Social Humanities Computing Laboratory at Tsinghua University, and the Language Computing and Internet Mining Research Group at Peking University. In industry, major e-commerce shopping sites such as Taobao, Tmall, Jingdong, Jindo and other social networking software such as Weibo and Twitter have applied sentiment computing technology to user review analysis. Through user review mining, manufacturers find problems with their products and improve them to enhance customer experience, thereby achieving the goal of increasing product sales.

Text sentiment analysis refers to the process of analyzing, processing and extracting subjective texts with emotional overtones using natural language processing and text mining techniques[22] . Text Sentiment Analysis[19] It is one of the important research areas in the field of affective computing[20] . The research scope of sentiment analysis includes opinion mining[22] , opinion mining[23] , opinion analysis[24] , Comment Mining[25] etc., is the process of analyzing, processing, summarizing and reasoning about the subjective information such as the opinions and attitudes or emotional tendencies of the expresser[26] The research is divided into Depending on the granularity of the research, it is mainly divided into Document- level, Sentence-level and Aspect-based [27].[27] The document-level sentiment analysis refers to the analysis of the document. Document-level sentiment analysis refers to the classification of the sentiment polarity of a document, i.e. positive and negative polarity. In addition to positive and negative polarity, the metric of a document or sentence may take different forms. For example, 'five-star' ratings are common in almost all assessment systems. In this case, sentiment polarity is divided into five levels. Sentence-level sentiment analysis is similar to document-level sentiment analysis in that it aims to extract the sentiment polarity of sentences. Traditionally, sentiment analysis has been carried out at a coarse-grained level for both chapter-level and sentence-level text. Coarse-grained sentiment analysis, however, only assumes that a text contains only a single sentiment, such as positive or negative, and is unable to identify sentiment for texts containing multiple aspects.

Research methods for sentiment classification are mainly divided into lexicon-based method and machine learning methods (e.g., Deep learning Deep learning method)[28] . Lexicon based sentiment analysis is mainly used to solve general sentiment analysis problems[28] It sometimes suffers from the low coverage of sentiment dictionaries[29] . To improve the accuracy of sentiment classification, lexicon-based approaches have recently been combined with machine learning methods. hailong et al.[30] and Ravi[31] found that machine learning methods were more accurate than lexical methods. Mudinas and Zhang et al.[32] improved the accuracy of sentiment analysis by combining lexicon and support vector machine (SVM) based methods.

Recently, deep learning methods have been playing an increasingly important role in natural language processing. most of the deep learning tasks in NLP have been targeted using word vector representation[33] . Word embedding (e.g. Word2Vec and GloVe) of continuous vector representations is a deep learning technique that converts words into meaningful vectors. Vector representations of words are useful in text classification, clustering and information retrieval. Due to the rise of neural networks, especially the successful application of attention and memory mechanisms to various natural language processing tasks, deep learning-based sentiment analysis methods can achieve superior results. With the development of deep learning, various neural networks are widely used to solve natural language processing (NLP) tasks, such as convolutional neural networks (CNNs)[34] , recurrent neural networks (RNNs)[35] [35], graph based neural networks (GNNs)[36] and attention mechanisms[37] etc. One of the advantages of these neural models is their ability to alleviate feature engineering problems. While non-neural NLP methods typically rely heavily on discrete manual features, neural methods typically use low-dimensional and dense vectors (also known as distributed representations) to implicitly represent the syntactic or semantic features of a language. These representations are learned in the context of a specific NLP task. Thus, neural methods make it easy to develop a variety of NLP systems.

Recently, there has been widespread interest in the field of natural language processing (NLP), where the use of language model pre-training methods has yielded good improvements on a number of NLP tasks. Pre-trained language models have been shown to be effective in improving many NLP tasks[38] . Despite the success of neural models in NLP tasks, the performance improvements may be less dramatic compared to the field of computer vision. A large body of recent work has shown that pre-trained models on large corpora can learn generic language representations, which can be beneficial for subsequent NLP tasks, avoiding the need to train new models from scratch. Models are not usually built and trained from scratch due to time constraints or hardware level limitations, and the use of pre-trained models allows for rapid implementation of models, which is why pre-trained models exist. Therefore, improving the accuracy of pre-trained word embeddings is important and plays a crucial role in sentiment classification methods. Zhang and Wallace[39] combined pre-trained Word2Vec and GloVe vectors in their deep learning model, but with reduced accuracy. Due to time constraints or hardware level limitations, in order to make better use of the pre-trained model and to better implement fine-grained sentiment analysis tasks, we conducted an in-depth study to improve the pre-trained model by incorporating more text features to achieve fine-grained aspect-level sentiment computation tasks.

3. Research on sentiment analysis methods for text-oriented data

This thesis is oriented towards the research of sentiment analysis methods for text data, in order to improve the accuracy, efficiency and cost reduction of sentiment analysis, with the application needs of massive text data as the traction, focusing on the challenges of text sentiment computation theory, pre-trained language model, aspect-level text classification, knowledge graph semantic understanding and other aspects, focusing on the research of text sentiment computation theory, text pre-trained language model, aspect-level text sentiment classification, The study focuses on the theory of text sentiment computing, pre-trained language models, aspect-level text sentiment classification, and semantic understanding of text based on knowledge graphs to achieve high accuracy and efficiency in text sentiment classification.

3.1. Hot research content

3.1.1. A computational theory of textual emotion.

Research Point 1 focuses on theories of textual affective computing. The development of affective computing research relies heavily on new advances in the study of human intelligence and emotion in the psychological and cognitive sciences. Textual affective computing focuses on the correspondence between affective states and textual information, providing clues to human affective states, finding features that can be extracted by computers and using models that can be used for affective classification, converting natural language text in text into a form that can be recognised and processed by computers, and using the affective information classification module to obtain computational results.

3.1.2. Text pre-training language models.

Pre-trained language models have developed a new paradigm for natural language processing: pre-training using large text corpora, fine-tuning small datasets for specific tasks, and reducing the difficulty of individual natural language processing tasks. A pre-trained language model is a sentence-level representation of word/word embeddings in context, containing rich syntactic and semantic information, which makes full use of a large-scale monolingual corpus, allowing the modelling of multiple meanings of a word and learning rich semantic knowledge from an unbounded large-scale monolingual corpus.

3.1.3. Aspect-level text sentiment classification.

Aspect-level Sentiment Analysis is more granular, with the task of extracting and summarizing people's opinions about an entity and the characteristics of the entity (the target). Aspect-level sentiment analysis consists of several subtasks, such as Aspect extraction, entity extraction and Aspect sentiment classification. Aspect-level sentiment classification takes into account both sentiment and topic information. Given a sentence and a topic feature, aspect-level sentiment classification can infer the sentiment polarity/proclivity of the sentence with respect to the topic feature. The modelling goal of aspect-level sentiment classification is difficult to correlate with the semantic relevance of contextual words in the context. To investigate how different contextual words, have different effects on the sentiment polarity of sentences at the target feature.

3.1.4. Knowledge graph-based text semantic understanding techniques.

Semantic understanding of text based on knowledge graph is to do all-round parsing of text from the knowledge dimensions of entity, concept and relationship to help provide the semantic knowledge required by the application. Firstly, the text is annotated with entity classes, then the entities are associated to the knowledge graph, and the corresponding information of the entities is obtained through the associated relationships as well as the knowledge graph; secondly, conceptualization is carried out to understand the knowledge behind the entities; finally, the relationships between the entities will be understood, including the attributes and sides of the entities. The semantic understanding of text that builds the knowledge graph will have technical features such as semantic disambiguation, computable reasoning and generalizable interpretation.

3.2. Key Issues Research

3.2.1. Textual affective computational theory reframes the problem.

How to carry out theoretical research on textual affective computing in conjunction with theoretical research on affective computing for textual data. Through the combination of computational science with psychological science and cognitive science, we study the characteristics of emotion during human-human interaction and human-computer interaction, design human-computer interaction environments with emotional feedback, study how to acquire and model emotional information, and how to better recognise emotion and understand and express emotion.

3.2.2. Accuracy issues in text pre-training language models.

How to better distinguish the different semantics of polysemous words/words and understand complex contexts for more accurate text representation, and how to improve the accuracy of pre-trained languages to improve sentiment classification calculations. How to use pre-trained models as benchmarks to improve existing models for more accurate text sentiment computation due to time constraints or hardware level limitations.

3.2.3. Aspect-level textual emotion object recognition issues.

A key step in text sentiment computing is sentiment object recognition. Sentiment objects include both fine-grained sentiment object recognition and implicit sentiment object recognition. In textual data, a sentence usually contains multiple objects, each holding different sentiment tendencies. Depending on the characteristics of the data, it is a challenge to construct a suitable model to solve the problem of object recognition and sentiment tendency determination.

3.2.4. Multidimensional semantic disambiguation issues for knowledge enhancement.

To address the problem of semantic disambiguation in the semantic understanding of knowledge graph text, how the text can be parsed from the knowledge dimensions of entities, concepts and relationships to do a full range of parsing to help provide the semantic knowledge required for applications and establish the semantic understanding of text for knowledge graph.

3.3. Commonly used experimental methods

Research methods used

The research for this thesis is based on literature research, experimental research, cross-sectional research and empirical findings.

3.3.1. Literature research method.

A method of obtaining information through a survey of the literature in order to gain a better understanding of the history and current status of the issue and to help define the research topic, so as to gain a comprehensive and correct understanding of the problem to be studied, in accordance with the content and purpose of the research project "Sentiment analysis methods for textual data".

3.3.2. The experimental research method.

The experimental research method is an important method of researching a subject. The experimental research method was first used in the natural sciences and has become its main research method. It is a method of drawing certain scientific conclusions by carrying out planned practice based on certain theories or hypotheses and testing them through experimental operations.

3.3.3. The cross-sectional approach.

A comprehensive approach to a subject using theories, methods and results from multiple disciplines, also known as the 'cross-research method'. The interdisciplinary approach refers to the interpenetration of disciplines, which simply means that the disciplines are no longer based on the knowledge of one discipline alone, but rather form a situation in which you have me and I have you.

3.3.4. Summary of experience method.

Based on one's accumulated technical experience combined with relevant theoretical research on the research object to share, from a large amount of research information to summaries the pattern, and through the research results to verify and enhance the hypothesis.

3.4. Technical lines

3.4.1. A computational theory of textual emotion.

Research point one proposes to expand the theory of textual emotion computing by building on existing theories of emotion computing. Text sentiment computing focuses on the correspondence between sentiment states and textual information, providing clues to human sentiment states, finding features that can be extracted by computers, and using models that can be used for sentiment classification to transform natural language text into a form that can be recognised and processed by computers, and using sentiment information classification modules to obtain computational results. Depending on the granularity of the text being processed, text sentiment computing can be broadly divided into three levels of research: word-level, sentence-level and chapter-level. The Internet (e.g. blogs and forums, and social service networks such as public reviews) generates a large number of user-participated reviews of valuable information such as people, events and products. These reviews express various emotional colours and tendencies, such as happiness, anger, sadness, joy and criticism, praise, etc. In terms of data sources, it is mainly based on online open data sets, such as the Weibo data set, Taobao, Jingdong, Twitter and other open source rich data.

3.4.2. Text pre-training language models.

A text pre-trained language model is a sentence-level representation of word/word embeddings in context, containing rich syntactic and semantic information, making full use of a large-scale monolingual corpus, which allows the modelling of multiple meanings of a word and the learning of rich semantic knowledge from an unbounded large-scale monolingual corpus. Sentiment analysis based on text pre-training performs sentiment classification by means of a text pre-training model and a sentiment classification model. The key model building steps are as follows.

3.4.3. Aspect-level text sentiment classification.

Aspect-level Sentiment Analysis is a more fine-grained sentiment analysis that extracts and summarises people's opinions about an entity and the characteristics of the entity (target), classifies the sentiment contained in different aspects of the text, and enables multidimensional sentiment understanding of the text. Aspect-based sentiment analysis consists of several subtasks, such as Aspect extraction, entity extraction and Aspect sentiment classification. Depending on the characteristics of the data, suitable models are constructed to solve object recognition and sentiment tendency determination problems.

3.4.4. Knowledge graph-based text semantic understanding techniques.

Semantic understanding of text based on knowledge graph is to do all-round parsing of text from the knowledge dimensions of entity, concept and relationship to help provide the semantic knowledge required by the application. Firstly, the text is annotated with entity classes, then the entities are associated to the knowledge graph, and the corresponding information of the entities is obtained through the associated relationships and the knowledge graph; secondly, conceptualization is carried out to understand the knowledge behind the entities; finally, the relationships between the entities will be understood, including the attributes and sides of the entities, etc. By introducing the corresponding knowledge graphs, knowledge-enhanced semantic analysis from multiple dimensions can be carried out for better semantic understanding, leading to a more prepared sentiment classification.

4. Conclusion

The rapid development of information technology has brought about a boom in development and a dramatic increase in the number of online platforms. The results obtained from sentiment analysis of the large amount of speech information on these platforms can be used to classify reviews on online platforms, analyse and recommend products and make consumption predictions, which have high commercial value. Traditional methods of text sentiment analysis are labour-intensive, but the limited coverage of manually extracted features and the irrational behaviour of humans can affect the correctness of the results, so traditional methods are not universally applicable. With the development of deep learning, pre-trained language models of text have emerged as a new paradigm for natural language processing. Pre-training is performed using a large text corpus to fine-tune small datasets for specific tasks and reduce the difficulty of individual natural language processing tasks. A pre- trained language model is a sentence-level representation of word/word embeddings in context, containing rich syntactic and semantic information, making full use of a large-scale monolingual corpus, allowing the modelling of multiple meanings of a word and learning rich semantic knowledge from an unbounded large-scale monolingual corpus.

For the study of sentiment analysis methods for text-oriented data, we propose to explore sentiment analysis based on pre-trained language, aspect-level sentiment analysis and knowledge graph-based understanding. In summary, the research work in this paper has good theoretical research significance, commercial value and application prospects. Text sentiment computing techniques are not only of important research value in theoretical studies such as learning models and algorithm optimisation, but also in engineering applications such as commodity evaluation, opinion retrieval, opinion detection and information prediction.

References

[1] W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms and applications: a survey," Ain Shams engineering journal, vol. 5, no. 4, pp. 1093-1113, 2014.

[2] M. Heidari and S. Rafatirad, "Using transfer learning approach to implement convolutional neural network model to recommend airline tickets by using online reviews," in 2020 15th International Workshop on Semantic and Social Media Adaptation and Personalization (SMA), pp. 1-6, 2020.

[3] M. Heidari and S. Rafatirad, "Semantic convolutional neural network model for safe business investment by using bert," in IEEE 2020 Seventh International Conference on Social Networks Analysis, Management and Security, SNAMS 2020, 2020.

[4] M. Heidari, J. H. J. Jones, and O. Uzuner, "Deep contextualized word embedding for text-based online user profiling to detect social bots on twitter," in IEEE 2020 International Conference on Data Mining Workshops (ICDMW), ICDMW 2020, 2020.

[5] M. Heidari and J. H. Jones, "Using bert to extract topicindependent sentiment features for social media bot detection," in 2020 11th IEEE Annual Ubiquitous Computing, Electronics Mobile Communication Conference (UEMCON), pp. 0542-0547, 2020.

[6] M. Heidari, S. Zad, B. Berlin, and S. Rafatirad, "Ontology creation model based on attention mechanism for a specific business domain, " in IEEE 2021 International IOT, Electronics and Mechatronics Conference, IEMTRONICS 2021, 2021.

[7] State Council. Notice of the State Council on the Issuance of the Action Outline for Promoting the Development of Big Data [OL]. China government website. Beijing: 2015.

[8] China Electronics Technology Standardization Institute, Working Group on Big Data Standards, National Information Technology Standardization Technical Committee. Big Data Standardization White Paper 2018 Edition, 2018-04.

[9] Pang B, Lee L, Vaithyanathan S. Thumbs up?: sentiment classification using machine learning techniques[C]. Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP). Philadelphia, PA, USA: Association for Computational Linguistics, 2002: 79-86.

[10] Liu B. Sentiment analysis: mining opinions, sentiments, and emotions[M]. Cambridge CB2 8BS UK: Cambridge University Press, 2015.

[11] E. Cambria, S. Poria, A. Hussain and B. Liu. Computational Intelligence for Affective Computing and Sentiment Analysis [J]. IEEE Computational Intelligence Magazine, 2019,14(2):16-17.

[12] Kim, Soo-Min & Hovy, Eduard. Determining the sentiment of opinions[C]//Proceeding of the 20th International Conference on Computational Linguistics. Association for Computational Linguistics, 2004:1367.

[13] Collobert R, Weston J, Bottou L, et al. Natural language processing (almost) from scratch[J]. Journal of machine learning research, 2011, 12(Aug): 2493-2537.

[14] Cambria E, Poria S, Gelbukh A, et al. Sentiment analysis is a big suitcase[J]. IEEE Intelligent Systems, 2017, 32(6): 74- 80.

[15] Chaturvedi I, Cambria E, Welsch R E, et al. Distinguishing between facts and opinions for sentiment analysis: Survey and challenges[J]. Information Fusion, 2018, 44: 65-77.

[16] Lo S L, Cambria E, Chiong R, et al. Multilingual sentiment analysis: from formal to informal and scarce resource languages[J]. Artificial Intelligence Review, 2017, 48(4): 499-527.

[17] Bernhard K , Suzana I , Mathias K , et al. Deep learning for affective computing: text-based emotion recognition in decision support[J]. Decision support systems, 2018, 115(NOV.):24-35.

[18] Cambria E . Affective Computing and Sentiment Analysis[J]. IEEE Intelligent Systems, 2016, 31(2):102-107.

[19] Zhao Yanyan, Qin Bing, Liu Ting. Text sentiment analysis[J]. Journal of Software, 2010, 21(8):1834-1848. ZHAO Yanyan, QIN Bing, LIU Ting. Sentiment analysis[J]. Journal of Software, 2010, 21(8):1834-1848.

[20] Saif H , He Y , Fernandez M , et al. Contextual semantics for sentiment analysis of Twitter [J]. Information Processing & Management, 2016, 52(1):5-19.

[21] Tang D , Wei F , Qin B , et al. Sentiment Embeddings with Applications to Sentiment Analysis[J]. IEEE Transactions on Knowledge and Data Engineering, 2015:1-1.

[22] Pang B , Lee L . Opinion Mining and Sentiment Analysis[J]. Foundations & Trends in Information Retrieval, 2008, 2(1-2):1-135.

[23] Liu, Bing. Sentiment Analysis and Opinion Mining[J]. Synthesis Lectures on Human Language Technologies, 2012, 5(1):1-167.

[24] Schouten K , Frasincar F . Survey on Aspect-Level Sentiment Analysis[J]. IEEE Transactions on Knowledge and Data Engineering, 2016, 28(3):813-830.

[25] Koyel Chakraborty, Siddhartha Bhattacharyya, Rajib Bag , et al. Comparative Sentiment Analysis on a Set of Movie Reviews Using Deep Learning Approach [C]// International Conference on Advanced Machine Learning Technologies & Applications. Springer, Cham, 2018.

[26] Giatsoglou M , Vozalis M G , Diamantaras K , et al. Sentiment analysis leveraging emotions and word embeddings[J]. Expert Systems with Application, 2017, 69(mar.):214-224.

[27] Wilson T, Wiebe J, Hoffmann P. Recognizing contextual polarity in phrase-level sentiment analysis. Int'l Journal of Computer Applications, 2005,7(5):347-354.

[28] Medhat W , Hassan A , Korashy H . Sentiment analysis algorithms and applications: a survey[J]. Ain Shams Engineering Journal, 2014, 5(4):1093-1113.

[29] Giatsoglou M , Vozalis M G , Diamantaras K , et al. Sentiment analysis leveraging emotions and word embeddings[J]. Expert Systems with Application, 2017, 69(mar.):214-224.

[30] Z. Hailong, G. Wenyan and J. Bo, "Machine Learning and Lexicon Based Methods for Sentiment Classification: A Survey," 2014 11th Web Information System and Application Conference, Tianjin, 2014, pp. 262-265.

[31] Kumar Ravi, Vadlamani Ravi, A survey on opinion mining and sentiment analysis: Tasks, approaches and applications, Knowledge-Based Systems, Volume 89 , 2015, Pages 14-46.

[32] Andrius Mudinas, Dell Zhang, and Mark Levene. 2012. Combining lexicon and learning based approaches for concept-level sentiment analysis. in Proceedings of the First International Workshop on Issues of Sentiment Discovery and Opinion Mining (WISDOM '12). Association for Computing Machinery, New York, NY, USA, Article 5, 1-8.

[33] Sanchez-Rada, J, Fernando, et al. Enhancing deep learning sentiment analysis with ensemble techniques in social applications[J]. Expert Systems with Application, 2017.

[34] Nal Kalchbrenner, Edward Grefenstette, and Phil Blunsom. A convolutional neural network for modelling sentences. in ACL, 2014.

[35] Kaitao Song, Xu Tan, Tao Qin, Jianfeng Lu, and Tie-Yan Liu. MASS: masked sequence to sequence pre-training for language generation. in ICML, volume 97 of Proceedings of Machine Learning Research, pages 5926-5936, 2019.

[36] Prajit Ramachandran, Peter J Liu, and Quoc Le. Unsupervised pretraining for sequence to sequence learning. in EMNLP, pages 383-391, 2017.

[37] Yu Sun, Shuohuan Wang, Yukun Li, Shikun Feng, Hao Tian, Hua Wu, and Haifeng Wang. ERNIE 2.0: A continual pretraining framework for language understanding. In AAAI, 2019.

[38] Andrew M Dai and Quoc V Le. Semi-supervised sequence learning. in NeurIPS, pages 3079-3087, 2015.

[39] Zhang, Y., & Wallace, B. (2015). A sensitivity analysis of (and practitioners' guide to) convolutional neural networks for sentence classification, arXiv, 1510.03820 v4 .

[40] Matthew Peters, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee, and Luke Zettlemoyer. 2018a. Deep contextualized word representations. in NAACL.

[41] Alec Radford, Karthik Narasimhan, Tim Salimans, and Ilya Sutskever. 2018. Improving language understanding with unsupervised learning. technical report, OpenAI. report, OpenAI.

[42] Jeremy Howard and Sebastian Ruder. 2018. universal language model fine-tuning for text classification. in ACL. association for Computational Linguistics.