

# Sentiment Analysis of Comment Texts

Subjects: **Computer Science, Artificial Intelligence**

Contributor: Baohua Su , Jun Peng

With information technology pushing the development of intelligent teaching environments, the online teaching platform emerges timely around the globe, and how to accurately evaluate the effect of the “any-time and anywhere” teacher–student interaction and learning has become one of the hotspots of today’s education research. Bullet chatting in online courses is one of the most important ways of interaction between teachers and students. The feedback from the students can help teachers improve their teaching methods, adjust teaching content, and schedule in time so as to improve the quality of their teaching.

sentiment analysis

attention mechanism

## 1. Introduction

In recent years, network technologies, such as the Internet, the Internet of things, and big data, have developed rapidly, and network platforms for e-commerce, social communication, and education are emerging timely. These platforms have not only enriched our daily life but also changed our ways of working, studying, and living. The sentiment comment texts on the network platform reflect people’s opinions on something. Thus, how to effectively use these opinions has become an important factor in improving service quality. In education, many countries have shifted their offline teaching to online teaching due to the global COVID-19 pandemic <sup>[1][2]</sup>. Compared with the traditional offline classroom, online education has the advantages of lower costs, flexible forms, and fewer geographical restrictions <sup>[3][4]</sup>. Its promotion and application increase the equity of higher education, realize knowledge sharing, improve the effectiveness and efficiency of decision-making, and make higher education more open <sup>[5]</sup>. In order to further evaluate the quality of teaching and strengthen the interaction between teachers and students, a large number of teaching platforms, such as China Universities MOOC and Tencent Classroom, have provided the bullet chatting function. The bullet chatting imbued with sentiment information plays an important role in the teaching process. Through students’ feedback, teachers can know what points students are weak in. School administrators can dynamically adjust the knowledge points, teaching plans, teaching objectives, and teaching staff structure of the courses based on the sentiment analysis of comment texts. Therefore, how to leverage useful information from comment text with sentiment information has become one of the hot research directions in natural language processing <sup>[6]</sup>.

Sentiment analysis is used to judge the sentiment polarity (positive, neutral, or negative) of reviews. Since Pang et al. studied the sentiment analysis of film reviews, sentiment analysis technology has been widely used in the business community <sup>[7]</sup>. As an emerging educational approach in the era of information technology, online courses have attracted many educators and learners around the world with their advantages of spanning time and space

and flexible learning methods. Comments, as the most direct way of interactive feedback in online courses, are of great significance in improving the quality of teaching, reducing the dropout rate, and promoting the sustainable development of online courses [8][9][10]. So, sentiment analysis is also very important in the field of education, but very few researchers do sentiment analysis in online course reviews, and even public data sets on this are very scarce.

There are three main methods of sentiment analysis: sentiment analysis based on sentiment dictionaries and rules, sentiment analysis based on traditional machine learning, and sentiment analysis based on deep learning [11]. Soe et al. further calculated sentiment scores to achieve the purpose of analyzing students' emotions through a part-of-speech tagging analyzer and vocabulary resources [12]. The second type of sentiment analysis method recognizes sentiment through constructing features artificially and using naïve Bayes, maximum entropy, and support vector machine and other classifiers.

With the development of deep learning and the improvement in text representation methods based on deep learning, many researchers began to study the application of deep learning to deal with text sentiment analysis. Represented by RNN, LSTM, and other classical neural networks, deep learning-based sentiment analysis methods can not only solve the shortcomings of traditional machine learning but also have significant classification effects. CNN can obtain the local information of a text, whereas recurrent neural networks such as LSTM can obtain the global information of a text. On the one hand, sequence-based neural networks such as LSTM have been restricted by the sequence length and computational memory. Attention mechanisms, on the other hand, could alleviate this problem since it allows modeling of the dependency output sequence without considering the distance between texts [13][14][15]. As a result, there are some sentiment analysis methods that use classical neural networks combined with an attention mechanism.

## 2. Sentiment Analysis of Comment Texts

Based on deep learning, there are two major categories of sentiment analysis models: graph-based models and sequence-based models.

The TextGCN model proposed by Yao et al. was the first time to use GCN in text classification (sentiment analysis) [16]. Two graphs were employed by that study as effective tools. The one named PMI was used to construct the relationship between words, and another named TF-IDF was used to construct the relationship between documents and words, and then the text category was obtained by the classifier. Then, Ragesh et al. [17] and Galke et al. [18] developed HeteGCN, which combined features of predictive text and TextGCN; It means the adjacency matrix was split into word documents and word submatrices, and the representations of different layers were fused as needed. Subsequently, HyperGAT was brought forward by Ding et al., from which an edge can connect multiple vertices [19]. So the text information was transformed into a hypergraph between nodes and edges, and the information between each layer was aggregated by dual attention. At last, tensorGCN was presented by Liu et al. [20]. This model constructed multiple graphs to describe semantic, syntactic, and contextual information and improved the effect of text classification through learning intra-graph propagation and inter-graph propagation.

Some studies have found that in recent years, most of the new methods for sentiment analysis (text classification) are based on GCN, while transformer-based sequence models are rare in the literature [18]. However, much empirical evidence shows that transformer-based sequence models outperform GCN-based methods. So here is a look at some sequence-based text classification methods. After obtaining the representation of each word, Kim embedded the word into CNN to obtain the sentiment polarity of the text [21]. Through the experimental results of a large number of data sets, he proved the ability of CNN on the task of text classification. After obtaining the text representation, Liu et al. used an RNN to classify the sentiment of the comment text [22]. Wang et al. proved that LSTM could achieve better experimental results than traditional RNNs in tweet sentiment analysis through experiments on tweet datasets [23]. After acquiring the word representation, the RNNs acquire the phrase representation and the sentence representation in order according to the syntactic structure. Huang et al. used a two-layer LSTM to classify the sentiment of tweets and believed that the sentiment polarity of the current tweet was largely related to the previous and subsequent tweets [24]. If the sentiment polarity is judged by the current tweet alone, the system would be deceived by its irony and other language expressions. Therefore, the hidden layer state of the current tweet should be input into a higher-level LSTM to obtain the current tweet representation containing context information and, finally, obtain the sentiment polarity distribution of the current tweet through the classifier. Yang et al. used the attention mechanism to aggregate word information to obtain sentence information, then they used the second layer attention mechanism to aggregate sentence information, in order to obtain the overall sentiment polarity in the discourse-level sentiment analysis, which fully proved the importance of an attention mechanism in sentiment analysis [25]. Vaswani et al. proposed the transformer model, which once again proved the importance of an attention mechanism in text classification [13]. Since the invention of BERT in 2018, there has been a lot of research on sentiment analysis based on BERT [26]. In Order to solve the negative effect of mask in BERT, XLNet uses an autoregressive language model instead of an autoencoding language model and introduces a double-stream self-attention mechanism and transformer-xl [27]. Compared with BERT, XLNet achieves better experimental results. ERNIE uses the same coding structure as BERT, but the author thinks that the random mask mechanism in BERT ignores the semantic relationship to some extent, so the original mask is split into three parts, the first part retains the original random mask, the second part masks the entity word as a whole. The last part is to mask the phrase as a whole. Compared with ERNIE, ERNIE 2.0 proposes three types of unsupervised tasks, which provide the model with a better representation ability of sentences, grammar, and semantics [28]. The performance and advantages of some methods on data sets are summarized in **Table 1**.

**Table 1.** Comparison with some methods.

Model	Data (acc)						Advantages
	SST-2	20NG	R8	R52	Ohsumed	MR	
Text GCN	-	0.863	0.970	0.935	0.683	0.767	A heterogeneous graph based on text and words is constructed, and the semi-supervised classification of text can be performed on GCN
HeteGCN	-	0.846	0.972	0.939	0.638	0.756	Reduce the complexity of TextGCN

Model	Data (acc)						Advantages
	SST-2	20NG	R8	R52	Ohsumed	MR	
HyperGAT	-	0.862	0.970	0.950	0.699	0.783	Capturing higher-order interactions between words while improving computational efficiency
TensorGCN	-	0.877	0.980	0.951	0.701	0.780	Rich multi-subgraph feature representation
LSTM	-	0.754	0.961	0.905	0.511	0.773	More effective way to process sequence data
BERT	0.928	-	-	-	-	-	The vector representation is rich, which overcomes the gradient problem of LSTM when solving long sequence data
ROBERTa	0.937	-	-	-	-	-	Raining models with larger corpora and sequences, dynamic MASK mechanism
XL-net	0.971	-	-	-	-	-	Autoregressive training method to overcome the shortcomings of bert
ernie	0.935	-	-	-	-	-	Taking advantages of The lexical, syntactic and knowledge information, large-scale text corpora and KGs to train an augmented language representation model

1. Tarkar, P. Impact of COVID-19 pandemic on education system. *Int. J. Adv. Sci. Technol.* 2020, 29, 3812–3814.
2. Zhou, L.; Wu, S.; Zhou, M.; Li, F. ‘School’s out, but class’ on’, the largest online education in the world today: Taking China’s practical exploration during The COVID-19 epidemic prevention and control as an example. *Best Evid. Chin. Edu.* 2020, 4, 501–519.  
“-“ indicates that the original paper was not tested on this data set.
3. Cao, R.; Xu, S.; Wang, X. Digitalization Leads the Future of Global Higher Education—Summary of the Main Session of the 2022 World MOOC and Online Education Conference. *China Educ. Informatiz.* 2023, 29, 82–95.
4. Wang, X.; Guo, S. Practice and Enlightenment of Online and Offline Integrated Teaching in Tsinghua University. *Mod. Educ. Technol.* 2022, 32, 106–112.
5. Global MOOC and Online Education Alliance. Trends, Stages and Changes of Digitalization of Higher Education: An Excerpt from Infinite Possibilities: Report on the Development of Digitalization of World Higher Education. *China Educ. Informatiz.* 2023, 29, 3–8.
6. Feng, C.; Li, H.; Zhao, H.; Xue, Y.; Tang, J. Attribute level sentiment analysis based on hierarchical attention mechanism and gate mechanism. *Chin. J. Inf. Technol.* 2021, 35, 128–136.
7. Pang, B.; Lee, L.; Vaithyanathan, S. Thumbs up? Sentiment classification using machine learning techniques. *arXiv* 2002, arXiv:cs/0205070.

8. Wang, L.; Hu, G.; Zhou, T. Semantic analysis of learners' sentiment tendencies on online MOOC education. *Sustainability* 2018, 10, 1921.
9. Mite-Baidal, K.; Delgado-Vera, C.; Solís-Avilés, E.; Espinoza, A.H.; Ortiz-Zambrano, J.; Varela-Tapia, E. Sentiment analysis in education domain: A systematic literature review. In *Proceedings of the Technologies and Innovation: 4th International Conference, CITI 2018, Guayaquil, Ecuador, 6–9 November 2018*; Springer International Publishing: Cham, Switzerland, 2018; pp. 285–297.
10. Pan, F.; Zhang, H.; Dong, J.; Shou, Z. Sentiment analysis of Chinese online course reviews based on efficient Transformer. *Comput. Sci.* 2021, 48, 264–269.
11. Xu, G.; Meng, Y.; Qiu, X.; Yu, Z.; Wu, X. Sentiment analysis of comment texts based on BiLSTM. *IEEE Access* 2019, 7, 51522–51532.
12. Soe, N.; Soe, P.T. Domain oriented aspect detection for student feedback system. In *Proceedings of the 2019 International Conference on Advanced Information Technologies (ICAIT), Yangon, Myanmar, 6–7 November 2019*; pp. 90–95.
13. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. In *Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017*; Volume 30.
14. Bahdanau, D.; Cho, K.; Bengio, Y. Neural machine translation by jointly learning to align and translate. *arXiv* 2014, arXiv:1409.0473.
15. Kim, Y.; Denton, C.; Hoang, L.; Rush, A.M. Structured attention networks. In *Proceedings of the International Conference on Learning Representations, Toulon, France, 24–26 April 2017*.
16. Yao, L.; Mao, C.; Luo, Y. Graph convolutional networks for text classification. *Proc. AAAI Conf. Artif. Intell.* 2019, 33, 7370–7377.
17. Ragesh, R.; Sellamanickam, S.; Iyer, A.; Bairi, R.; Lingam, V. Hetegcn: Heterogeneous graph convolutional networks for text classification. In *Proceedings of the 14th ACM International Conference on Web Search and Data Mining, Virtual Event, Israel, 8–12 March 2021*; pp. 860–868.
18. Galke, L.; Scherp, A. Bag-of-words vs. graph vs. sequence in text classification: Questioning the necessity of text-graphs and the surprising strength of a wide MLP. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics, Dublin, Ireland, 22–27 May 2022*; pp. 4038–4051.
19. Ding, K.; Wang, J.; Li, J.; Li, D.; Liu, H. Be more with less: Hypergraph attention networks for inductive text classification. *arXiv* 2020, arXiv:2011.00387.
20. Liu, X.; You, X.; Zhang, X.; Wu, J.; Lv, P. Tensor graph convolutional networks for text classification. *Proc. AAAI Conf. Artif. Intell.* 2020, 34, 8409–8416.

21. Kim, Y. Convolutional neural networks for sentence classification. arXiv 2014, arXiv:1408.5882.
22. Liu, P.; Qiu, X.; Huang, X. Recurrent neural network for text classification with multi-task learning. In Proceedings of the 25th International Joint Conference on Artificial Intelligence, New York, NY, USA, 9–15 July 2016; pp. 2873–2879.
23. Wang, X.; Liu, Y.; Sun, C.J.; Wang, B.; Wang, X. Predicting polarities of tweets by composing word embeddings with long short-term memory. In Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers), Beijing, China, 26–31 July 2015; pp. 1343–1353.
24. Huang, M.; Cao, Y.; Dong, C. Modeling rich contexts for sentiment classification with lstm. arXiv 2016, arXiv:1605.01478.
25. Yang, Z.; Yang, D.; Dyer, C.; He, X.; Smola, A.; Hovy, E. Hierarchical attention networks for document classification. In Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, San Diego, CA, USA, 12–17 June 2016; pp. 1480–1489.
26. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv 2018, arXiv:1810.04805.
27. Yang, Z.; Dai, Z.; Yang, Y.; Carbonell, J.; Salakhutdinov, R.R.; Le, Q.V. Generalized autoregressive pretraining for language understanding. arXiv 2019, arXiv:1906.08237.
28. Sun, Y.; Wang, S.; Li, Y.; Feng, S.; Tian, H.; Wu, H.; Wang, H. Ernie 2.0: A continual pre-training framework for language understanding. Proc. AAAI Conf. Artif. Intell. 2020, 34, 8968–8975.

---

Retrieved from <https://encyclopedia.pub/entry/history/show/117669>