REVIEW ARTICLE

A Comprehensive Review on Techniques in Sentiment Analysis for Improving Teaching and Learning through Students' Feedback

Sowmiya Baskar[1,*]

[1] Department of Computer Science, S.I.V.E.T. College, Chennai 600073, Tamil Nadu, India

Abstract

Getting feedback from the students in education is the key to improving the learning experience in education, but reading through hundreds of feedback forms can be overwhelming. Sentiment Analysis (SA), which is a NLP (Natural Language Processing) technique, comes in interprets the emotions and opinions behind their feedback. This review explores how various technologies like machine learning and NLP are being used to understand student opinions about teaching quality, course materials, assignments, exams, instructional behavior and overall learning experience. Sentiment analysis helps educators understand student concerns, thereby improving the learning experience and promoting a student-centered learning environment. The challenges in these technologies are discussed with future directions. This review would serve as a guide to researchers in the same domain.

Keywords: sentimental analysis, student feedback, machine learning, natural language processing, bias.

Submitted: 02 September 2025
Accepted: 08 September 2025
Published: 11 October 2025

Vol. 1, **No.** 1, 2025.
10.62762/NGCST.2025.927566

***Corresponding author:**
Sowmiya Baskar
sowmiyabaskar@gmail.com

1 Introduction

Teacher evaluations through students' feedback provide a support system for improving teaching experience. Traditional evaluation lacks in capturing nuances from the feedback. Sentimental analysis plays a vital role in capturing the strengths and weaknesses from students' feedback. It directly supports teachers for further development. Sentiment mining may also be termed as opinion mining. The opinions are usually gathered at the end of every semester from students. Manually understanding students' feedback requires pedagogical expertise. Moreover, the labeling process is time-consuming. Advanced technologies like machine learning, deep learning and transformers etc. will help analyse the feedback of the students. Machine learning models like Support Vector Machine (SVM) effectively classify sentiments from the feedback of the students. These models have evaluation metrics like specificity and sensitivity, which offer confidence in the reliability of the model [1]. Automated sentiment analysis works across different levels like document, sentence and aspect while addressing the challenges like sarcasm handling [2].

For better management of the nuances of real-world student commands, researchers combine

Citation
Baskar, S (2025). A Comprehensive Review on Techniques in Sentiment Analysis for Improving Teaching and Learning through Students' Feedback. *Next-Generation Computing Systems and Technologies*, 1(1), 11–17.

lexicon-based methods and machine learning models as a hybrid sentiment classification model [3]. There are certain drawbacks in shallow ML models, as identified by researchers with bag-of–words and n-gram models. They struggle with capturing word order and semantics. To address this hybrid deep-learning model combining word attention, attention models and some linguistic rule was found to perform better by understanding nuances in feedback [4]. Feedback reports in Static PDF form have open-ended questions. Researchers worked on a method named SETSum, which applies sentiment analysis, summarization and visualization. SETSum also provided actionable feedback [5]. The rest of the paper will detail the techniques, relevant case studies, challenges and future directions.

The remaining sections of the paper are organized as follows. Section 2 explores the techniques of sentimental analysis. Section 3 discusses addressed challenges and limitations in sentimental analysis. Section 4 discusses the case studies and implementation. Section 5 explores ethical considerations. Section 6 explores the future directions and section 7 concludes the research.

2 Techniques in Sentiment Analysis

This section elaborates on the techniques used in sentimental analysis. Sentiment is extracted from student feedback, which is critical for developing more actionable insights. It actually started with simple rule-based systems; Progressing towards machine learning, then to advanced deep learning models, then attention mechanisms, the use of hybrid models and Transformer-based models. A notable approach is Aspect-Based Sentiment Analysis (ABSA) used nowadays. Multimodal research is an emerging one. Shaik et al. [2] in their comprehensive survey plot the succession from machine learning models like SVM, Naive Bayes, decision trees to deep learning architectures like LSTM and CNN and to transformer-based models like BERT. Misuraca et al. [4] developed a hybrid deep-learning model (DTLP) that integrates CNN and BiLSTM that capture both local features and long-term dependencies; the Attention mechanism focuses on the sentiment parts/portions of text; Word embeddings for richer semantic representation and Linguistic rules to handle the shifts (negation and sentiment polarity). Being tested on real-world feedback datasets and the hybrid DTLP model outperformed baseline models like BoW (Bag of Words) and CNN (Convolutional Neural Networks).

Hu et al. [5] developed SETSum, a system that analyzes student feedback using sentiment analysis, aspect extraction, and text summarization. They designed a visual dashboard to display summarized feedback for swift understanding. They have tested it with 10 professors, who found it faster and more useful than traditional SET reports. Hajrizi et al. [6] ABSA addresses a key limitation in traditional sentiment analysis, like the inability to detect what exactly students are commenting on. For instance, in the sentence: "The professor was great, but the homework was overwhelming," In the above feedback, a generic sentiment classifier may struggle to interpret the tone. The instructor is about to carry a positive tone and the assignment carries a negative tone. ABSA identifies both aspect terms (“professor”, “homework”) and assigns sentiment polarity to each. “The lectures were interesting (positive), but the assignments were confusing (negative).”.

Koufakou’s [7] work aims to exhibit how the latest NLP models can efficiently extract sentiment and topic information from the students’ feedback. They examine how transformer-based deep learning models like RoBERTa and BERT outperform classical ML in extracting sentiment from real student opinions. They explored hyperparameter optimization strategies for enhancing model performance.

Yang et al. [8] used GANs (CatGAN and SentiGAN) to synthetically generate student feedback, balancing datasets before training. They have trained both classical and deep learning classifiers on the new balanced data and attained 2.8%–9.3% improvements in F1-scores. Malik et al. [15] proposed a hybrid deep learning model utilizing stacked text embeddings and a multi-head attention mechanism to capture rich context. They have combined multiple deep learning classifiers into an ensemble for robust performance and achieved 95% accuracy, 97% recall and 96% F1-score. Figure 1 depicts a word cloud of techniques used in sentiment analysis.

Baragash et al. [9], through their research, bring to light that lexicon-based tools are rarely used and advanced techniques like deep learning and Aspect-Based Sentiment Analysis (ABSA) are insufficiently examined. Zhou et al. [10] provide a review of sentiment analysis applications in education. They focus on lexicon and machine learning approaches in formative assessments and feedback. Students’ reflections are focused, where student

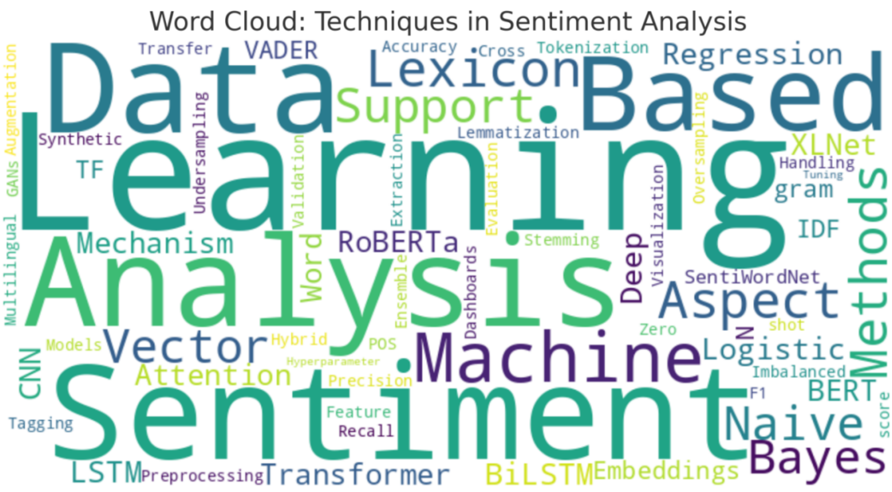

Figure 1. Word cloud of techniques in sentimental analysis.

satisfaction, dropout prediction and performance are assessed and visualized. Lexical tools like VADER, Textblob, SentiStrength, and SentiWordNet are used along with machine learning algorithms like Naïve Bayes, Support Vector Machine, K-Nearest Neighbor, etc. They insist on customized learning or self learning as a future tool. Aryal [11] explored the use of academic tools like Mentimeter, Kahoot and Padlet for gathering student feedback and examined sentimental analysis of comments. They also proposed a model for sentimental analysis which starts with data collection, then data is preprocessed, features are extracted and finally the model is classified and returns as positive or negative comments.

Pooja et al. [12] conducted a review of sentiment analysis to meet quality goals in education. The study highlights performance prediction using sentiments in social media platforms like Twitter and Facebook with classifiers like SVM and Naive Bayes classifiers. SVM is said to outperform in handling non-linear feedback data. Liu et al. [13] utilized multimodal feature fusion for sentiment analysis. Text data is enhanced by adding semantic context. Unimodal feature extraction is done to extract text and image features. Multimodal feature fusion is done where extracted text and image features are fused with the use of attention mechanisms. In the final layer, the classifier labels the sentiment.

El Maazouzi et al. [14] executed detection of emotional and cognitive states in multimodal data through NLP techniques. They effectively captured fatigue level, facial emotion and sentimental emotions in text from the learner. Public datasets like DAiSEE, AffectSet are used for fatigue detection and Facial emotion detection. The MOOC Review dataset is used for text data sentimental analysis. Different algorithms like CNN-LSTM, EmoNet and LSTM, and a fine-tuned BERT model are used. PERCLOS (percentage of eye closure) and yawning frequency are used as indicators for fatigue detection. All three outputs are combined using deep fusion methodology to predict the final output. Results indicated higher accuracy for fatigue detection and sentiment analysis.

3 Addressed Challenges and Remaining Limitations

This section explores the challenges and limitations addressed by various researchers. When it comes to interpreting sarcasm or ironic expressions, sentiment analysis models struggle. For example, "The lectures were so exciting that I slept through them." Considering this statement, “so exciting” makes positive, “I slept through them” contradicts the excitement and creates a sarcastic tone and meaning the true intent is negative. So, the final polarity is negative. However, most standard classifiers, lacking sarcasm awareness, may misclassify it as positive due to the presence of words like "exciting," thereby failing to capture the true negative sentiment. Also lack of data scarcity in education domain makes it hard to train and compare models consistently. Imbalanced datasets may also lead to biased decisions in feedback. Educational institutions receive feedback in multiple languages. But most of the SA systems are English-centric, which makes them less suitable for applicability [2].

Traditional models result in word embedding ambiguity; they generate a single vector for each word, often failing to differentiate multiple meanings. Some word models can’t tell if a word is positive or negative.

So, they treat opposite words like "good" and "bad" as if they mean the same. This makes it hard to understand the true feeling behind the text. This problem is known as sentiment-blind embeddings. Simple models like Bag-of-Words and n-grams don't consider the order of words, which may lead to confusion. For example, they may treat "The course was difficult" and "No course was difficult" as the same, even though the meanings are opposite. To fix this, one study used a more advanced method called CNN-BiLSTM with attention, which helps the model understand the correct meaning by looking at word order and context. The model used is quite complex because it combines word embeddings, CNN, BiLSTM, attention layers and some language rules. This makes it powerful but also demands more computing power and time to fine-tune [4].

Traditional machine learning models like SVM, Naive Bayes and ANN depend a lot on manually chosen features such as TF-IDF and n-grams. These methods often miss the deeper meaning or context in feedback. To handle this, the authors used a hybrid approach that combines sentiment dictionaries (lexicons) with machine learning, helping the model understand the feedback more accurately. ML classifiers struggle when trained on small datasets or dataset is unbalanced. This problem can be addressed using feature selection techniques. Instead of lexicon-only approaches, lexicon approaches with machine learning enhances adaptability and reduces misclassification [3]. The authors address the lack of a large labeled dataset as a hindrance for model training and evaluation. They also mention aspect extraction as a limitation, where understanding context and inter-aspect relationships is a difficulty; preprocessing variability makes the result comparison tedious [6].

Text preprocessing may sometimes be challenging due to typos, abbreviations or short texts, due to which model performance will be degraded. Though transformers like BERT, RoBERTa achieve high accuracy, they demand more training time and computational resources than traditional machine learning models [7]. A highly imbalanced dataset may affect model training. In such a case, GAN-based text generators can be used for creating synthetic minority-class samples, which help rebalance the dataset and improve the performance of the model. But synthesised datasets may be of low quality and fail to enrich minority class samples. This problem can be addressed by focusing on diverse sample generation with category-aware GANs. They enhance reliability. The model's performance from synthetic data may vary across datasets and models, like ML and DL. DL shows an improved F1-score than classical ML models with multiple datasets [8].

Other limitations include inadequate handling of figurative language and computational complexity [15]. Imbalanced classes affect the performance of the classifier, which may be solved using data augmentation and class-weighting strategies during training. On the other hand, deep learning models are powerful, yet they are black boxes. In order to reveal keywords that influence sentiment decisions, explainability techniques like attention heatmaps or SHAP values should be used [16]. The majority of the studies report lack of multilingual support or non-English MOOC affects global applicability [17].

4 Case Studies and Implementations

This section presents two case studies from Park et al. [18] and Koufakou [7].

4.1 Case Study 1: Sentiment Analysis in Educational Contexts

Park et al. [18] collected and analyzed preservice teachers' reflective journals, who participated in an educational course in Miswestern U.S. These data were preprocessed by removing titles, headers, footers and reflection prompts. They were converted into paragraphs using Pandas dataframe. Three popular LLMS: GPT-4, Gemini, and BERT were used to analyze the tone and emotions on a scale of 0-2. Each of the LLM invoked API with well-designed prompts. Each model analyzed five times and the results indicate that GPT-4 yielded a tone score of 1.5/1.5, Gemini with 1.5/1.0 in the fifth trail. BERT produced 84.2% optimism score and a 54.9% negative score. They conclude that LLMs can help the teachers grow and reflect, but they must be careful in understanding the subtle emotions and complex circumstances in teaching.

4.2 Case Study 2: Analyzing Student Course Feedback with Deep Learning

Koufakou [7] applied deep learning for opinion mining and topic classification for online course reviews. They collected data from publicly available course reviews (online). For opinion mining, labels are decided and preprocessing for the text is carried out. They used models like BERT, RoBERTa, XLNet (transformer-based) and baseline methods like

SVM and Logistic Regression for sentiment analysis. Fine-tuning is carried out and hyperparameter optimization is done. Model is evaluated with metrics like accuracy, precision and F1-score. Experimental results indicate RoBERTa with 95.5% accuracy and 84.7% F1-score. They conclude that the institutions can combine transformer models for sentiment analysis and traditional Machine Learning models (like SVMs) for the course review to analyze student feedback effectively.

5 Ethical Considerations

This section discusses the ethical considerations for Sentiment Analysis in Education. The general ethical considerations include the following. When the feedback is collected from students, proper information consent should be obtained, which ensures their data privacy. Models may produce biases due to gender, culture or race; this issue can be sorted using careful selection of data and a diverse dataset. Explainable AI (XAI) matters, because people should know why a particular decision has been taken by the system. Sentiment Analysis plays a key role, where if a system makes something wrong or sensitive decisions, it may embarrass people, their sentiments. In that case, XAI is not optional. Several researchers insisted that ethical considerations be imposed, as discussed in the rest of the section.

Barker et al. [19] highlighted that emotion detection faces ethical challenges. A Model can make biased decisions due to race, gender, ethnicity, or neurodiversity, which may impact a cluster of society. To address this biased decision-making, authors recommend diverse use of the dataset and creating an audit to ensure balanced performance. Also, misinterpretations may cause psychological impact on people. To tackle this issue, informed consent is to be implemented. Research data collected from people may be manipulated or extended for some purpose. For instance, though the original intention may help educators understand and engage students, the surveillance at all times may create potential misuse. Though the technology is created for a good purpose, it may mislead if it is uncontrollable. That's why XAI is essential.

Mohammad [20] addressed several ethical considerations in Emotion recognition and Sentiment Analysis. The emotions may be complex, in some cases nuances may be missed, leading to misunderstanding. Even though data is collected with consent, people may be unaware of it. Data transparency and protecting data are crucial in that case. Again, the model may produce biased decisions, due to lack of availability of diverse data. Emotion detection systems take decisions, which are sometimes diffcult to audit. In such a case, Explainability is essential. Moreover, sentiment analysis decisions should not hurt people, their emotions, etc. Fair use of emotion recognition will ensure a reputation for the system. Another potential ethical consideration is dual-use or misuse. This technology can be repurposed, but should follow strict ethical guidelines in order to avoid misuse.

6 Future Directions

Though recent advancements in SA have significantly improved the understanding of student opinions or feedback still there is room for improvement. One key direction is to develop multilingual and culturally adaptive models that can accurately interpret feedback across different languages and educational settings. Sarcasm, irony and figurative language are common in informal student expressions. More refined models are needed for detecting them.

Adding Aspect-Based Sentiment Analysis (ABSA) helps us to figure out what exactly students liked or disliked and not just the overall mood. For example, instead of saying a course is "good" or "bad," ABSA can show if students enjoyed the teaching style, found the materials helpful, or thought the assessments were too hard. This gives a more detailed and useful picture of their feedback. Real-time sentiment analysis system is a promising area, enabling educators to respond to student concerns during the course rather than after it ends. Explainable AI (XAI) models can improve transparency and trust, and provide clear justifications for their predictions. This could be one of the future directions. Students' feedback which is noisy, short, and grammatically inconsistent feedback remains a technical challenge that calls for more robust preprocessing and modeling techniques. Creating larger and well-balanced labeled datasets will help train models more effectively and allow for fair comparisons across systems. Furthermore, incorporating transfer learning and domain adaptation techniques could make models more flexible across subjects and institutions without needing complete retraining. In the future, it's essential to address privacy and ethical concerns, making sure that sentiment analysis tools respect student confidentiality and are used in a responsible, transparent manner.

7 Conclusion

Sentimental analysis has evolved from simple rule-based methods to advanced deep learning approaches, significantly improving the way student feedback is interpreted. Models like CNN-BiLSTM with attention and hybrid ML approaches have enhanced sentiment detection by capturing contextual and emotional nuances. Despite progress, challenges like sarcasm, short-text ambiguity, and domain-specific language still affect model accuracy. Several studies addressed class imbalance using data augmentation techniques, while others improved interpretability using visualization tools. Aspect-based Sentiment Analysis emerged as a promising direction to understand feedback on specific elements like teaching style or materials. Ethical concerns around student privacy and consent must be considered in the real-world deployment of sentiment tools. Future work should focus on building multilingual, domain-specific datasets and refining models to handle informal or noisy input. Ultimately, well-designed sentiment analysis systems can support educators in enhancing course quality and student engagement. This study explores how sentiment analysis helps improve the overall learning process in education and shapes the future of education. This review will benefit researchers in the same field.

Data Availability Statement

Not applicable.

Funding

This work was supported without any funding.

Conflicts of Interest

The author declares no conflicts of interest.

Ethical Approval and Consent to Participate

Not applicable.

References

[1] Pooja, & Bhalla, R. (2022). A review paper on the role of sentiment analysis in quality education. *SN Computer Science*, *3*(6), 469. [CrossRef]

[2] Shaik, S. M., Anbarasi, M., & Jayakumar, V. (2023). Sentiment analysis and opinion mining on educational data: A survey. *arXiv preprint arXiv:2302.04359*. [CrossRef]

[3] Nasim, Z., Rajput, Q., & Haider, S. (2017, July). Sentiment analysis of student feedback using machine learning and lexicon based approaches. In *2017 international conference on research and innovation in information systems (ICRIIS)* (pp. 1-6). IEEE. [CrossRef]

[4] Misuraca, M., Viteritti, A., & D'Ascenzo, F. (2023). Student's feedback analysis model using deep learning-based method and linguistic knowledge for intelligent educational systems. *Soft Computing, 27*, 12507–12519. [CrossRef]

[5] Hu, Y., Zhang, S., Sathy, V., Panter, A. T., & Bansal, M. (2022). SETSum: Summarization and visualization of student evaluations of teaching. In *Proceedings of NAACL: HLT Demonstrations* (pp. 71–89). Association for Computational Linguistics. [CrossRef]

[6] Hajrizi, R., & Nuçi, K. P. (2020). Aspect-based sentiment analysis in education domain. *arXiv preprint arXiv:2010.01429*. [CrossRef]

[7] Koufakou, A. (2023). Deep Learning for Opinion Mining and Topic Classification of Course Reviews. *arXiv*. [CrossRef]

[8] Yang, R., & Edalati, M. (2021). Using GAN-based models to sentiment analysis on imbalanced datasets in education domain. *arXiv*. [CrossRef]

[9] Baragash, R., & Aldowah, H. (2021). Sentiment analysis in higher education: A systematic mapping review. *Journal of Physics: Conference Series, 1860*, Article 012002. [CrossRef]

[10] Zhou, J., & Ye, J. (2020). Sentiment analysis in education research: A review of journal publications. *Interactive Learning Environments*, 1–13. [CrossRef]

[11] Aryal, H. (2021). A literature survey on student feedback assessment tools and their usage in sentiment analysis. *arXiv*. [CrossRef]

[12] Pooja, P., & Bhalla, R. (2022). A review paper on the role of sentiment analysis in quality education. *SN Computer Science*, *3*(6), Article 469. [CrossRef]

[13] Liu, Z., Cai, L., Yang, W., & Liu, J. (2024). Sentiment analysis based on text information enhancement and multimodal feature fusion. *Pattern Recognition, 156*, Article 110847. [CrossRef]

[14] El Maazouzi, Q., & Retbi, A. (2025). Multimodal detection of emotional and cognitive states in e-learning through deep fusion of visual and textual data with NLP. *Computers, 14*(8), 314. [CrossRef]

[15] Malik, S. Z., Iqbal, K., Sharif, M., Shah, Y. A., Khalil, A., Irfan, M. A., & Rosak-Szyrocka, J. (2024). Attention-aware with stacked embedding for sentiment analysis of student feedback through deep learning techniques. *PeerJ Computer Science, 10*, e2283. [CrossRef]

[16] Esparza, J. F., Agrawal, A., & Khan, M. (2022). Enhancing sentiment analysis on educational feedback using hybrid deep learning models. *arXiv preprint arXiv:2301.08826*. [CrossRef]

[17] Dalipi, F., Zdravkova, K., & Ahlgren, F. (2021).

Sentiment analysis of students feedback in MOOCs: A systematic literature review. *Frontiers in Artificial Intelligence*, *4*, 728708. [CrossRef]

[18] Park, Y., & Hong, Y. (2024, September). Sentiment analysis of preservice teachers' reflections using a large language model. In *2024 6th International Workshop on Artificial Intelligence and Education (WAIE)* (pp. 61-65). IEEE. [CrossRef]

[19] Barker, D., Tippireddy, M. K. R., Farhan, A., & Ahmed, B. (2025). Ethical Considerations in Emotion Recognition Research. *Psychology International*, *7*(2), 43. [CrossRef]

[20] Mohammad, S. M. (2022). Ethics Sheet for Automatic Emotion Recognition and Sentiment Analysis. *Computational Linguistics*, *48*(2), 239–278. [CrossRef]

Sowmiya Baskar received the Ph.D. degree in Computer Science from Bharathidasan University, Trichy, India in 2024. She secured first place in M.Sc., Computer Science at Bharathidasan University, Trichy, India, and holds a Ph.D. in Computer Science from the same university. With expertise in machine learning, deep learning, and image processing, she has published research papers at international conferences, UGC-CARE-listed and Scopus-indexed journals. Dr. Sowmiya is also a respected reviewer in Scopus indexed journals. She has been an invited speaker and has chaired sessions at international conferences. Passionate about teaching and research, she mentors students to excel in publishing and presenting their work globally. (Email: sowmiyabaskar@gmail.com)