Journal of Network Intelligence ©2025 ISSN 2414-8105 (Online)
Taiwan Ubiquitous Information Volume **10**, Number **4**, November **2025**

Sentiment Analysis of Student Comment Text on Online Education Platforms Using a Hybrid GRU-FCM Model

June Xu*

International School
Zibo Vocational Institute, Zibo 255300, China
junexu1204@outlook.com, 12040@zbvc.edu.cn

*Corresponding author: June Xu
Received June 25, 2024, revised January 15, 2025, accepted July 5, 2025.

Abstract. *In this study, the author presents a novel approach for analyzing the sentiment of student feedback on e-learning platforms using a hybrid model combining Convolutional Neural Networks (CNN) and Fuzzy C-Means (FCM) clustering. The CNN extracts rich, high-dimensional features from student comments, capturing intricate patterns and contextual nuances. These features are then processed by the FCM algorithm, which assigns fuzzy membership values to each comment, allowing for more nuanced sentiment classification. By integrating CNN's robust feature extraction capabilities with FCM's ability to handle uncertainty and ambiguity, our model achieves superior performance in sentiment analysis. Experimental results on a dataset collected from central Chinese Massive Open Online Course (MOOC) platforms demonstrate the effectiveness of our approach, showing significant improvements in accuracy and precision compared to traditional methods. This research contributes to enhancing the understanding of student sentiments, thereby aiding educators and platform developers in improving the quality of e-learning experiences.*
Keywords: Sentiment Analysis, E-Learning, Fuzzy C-Means Clustering, Hybrid Model, MOOC Platforms, Educational Technology, Text Classification.

1. **Introduction.** Advancements in technologies like the Internet, IoT, and big data have transformed education [1], with online learning becoming a key alternative to traditional methods due to its accessibility, affordability, and scalability [2][3]. Massive Open Online Courses (MOOCs) now enable millions of students worldwide to engage in interactive and flexible learning environments [4]. In this context, sentiment analysis is essential for understanding student feedback, helping educators assess teaching effectiveness, address challenges, and enhance e-learning experiences. Sentiment analysis, while effective in many domains, often struggles with issues such as ambiguous textual expressions, imbalanced datasets, and privacy concerns. These limitations hinder educators from fully utilizing feedback to enhance course quality and student engagement.

Major e-learning platforms like China's MOOC have implemented features such as bullet-pointed chats to evaluate teaching quality and increase engagement between educators and students. These emotive comments reflect individual perspectives, and effectively utilizing this feedback is crucial for improving service quality. These bullet-pointed interactions play a significant role in the educational process by helping teachers identify student weaknesses, which in turn aids in adjusting course content, teaching strategies, learning objectives, and faculty assignments [5]. Sentiment analysis is a crucial technique for assessing the emotional tone of textual data, categorizing sentiments into positive, neutral, or negative. Initially used mainly in business contexts, sentiment analysis has been applied to various fields such as healthcare, politics, finance, and education. For example, researchers have used sentiment analysis to predict healthcare outcomes from patient feedback, assess political views from social media, and forecast stock market trends based on investor sentiment [6]. The wide application of sentiment analysis underscores its potential for enhancing data-driven decision-making in research and practical settings. However, the rapid growth of this technology raises ethical concerns, including privacy, bias, and fairness. If the perspectives of marginalized groups are not adequately represented, the resulting analysis may be biased and misleading [7]. The inherent subjectivity in determining the sentiment of textual expressions remains a significant challenge. Thus, there is a need for robust, ethically sound sentiment analysis methodologies that address these issues and can more accurately interpret the nuances of human emotions and language. With thoughtful advancements, sentiment analysis can provide valuable insights from diverse narratives and perspectives.

Researchers have primarily employed three approaches to sentiment analysis: dictionary and rule-based methods, traditional machine learning-based methods, and deep learning-based methods [8]. Despite its widespread use in business, sentiment analysis of online course reviews remains relatively underexplored, and publicly available datasets on this topic are scarce [9]. Various methods have been used to calculate sentiment scores in online education contexts. Traditional machine learning methods involve feature construction and classifiers like Naive Bayes, maximum entropy, and support vector machines (SVM) [10]. However, the accuracy of these methods depends heavily on the quality of feature selection and classifier performance. Deep learning techniques, such as CNN, have been developed to overcome the limitations of traditional methods. These advanced neural network architectures, often enhanced with attention mechanisms, have improved performance in capturing long-range dependencies and contextual information in text [11]. Sentiment analysis in education is crucial for enhancing teaching quality, reducing dropout rates, and promoting sustainable educational development. This study proposes a novel approach that combines FCM clustering with CNN to enhance sentiment analysis accuracy in evaluating online course feedback. Feature engineering in this context faces challenges related to high-dimensional data, feature sparsity, and domain adaptation [12]. Therefore, innovative techniques are necessary to address these challenges and improve the precision of sentiment analysis [13].

2. **Related Work.** Sentiment analysis has become a pivotal research area in natural language processing (NLP), classifying textual data into positive, negative, or neutral sentiments. Traditional approaches have evolved from basic machine learning techniques to sophisticated deep learning models, enhancing the accuracy and efficiency of sentiment classification. Early sentiment analysis efforts utilized machine learning classifiers such as Naive Bayes, SVM, and maximum entropy models. These methods required extensive feature engineering, including features based on term frequency-inverse document frequency (TF-IDF) and other lexical indicators. While effective, these models heavily depended on the quality of feature extraction and could not capture complex linguistic nuances.

The advent of deep learning revolutionized sentiment analysis by introducing models capable of automatically learning feature representations from raw text data. CNN and Recurrent Neural Networks (RNNs), particularly LSTM networks, became prominent due to their ability to capture contextual information and sequential dependencies in text. For instance, CNNs are adept at identifying local patterns in text, making them suitable for tasks such as sentence classification [14]. LSTM networks, on the other hand, excel at capturing long-range dependencies and contextual nuances, which are crucial for understanding the sentiment in lengthy texts [15]. Attention mechanisms and Transformer-based models have further advanced sentiment analysis. The Transformer architecture, introduced by Chanaa et al., relies on self-attention mechanisms to model dependencies between words, regardless of their distance in the text [16]. This has led to the development powerful models such as Bidirectional Encoder Representations from Transformers (BERT) [17], which can be fine-tuned for various NLP tasks, including sentiment analysis. BERT's bidirectional nature allows it to consider the context from both directions, enhancing its understanding of the text. Extensions of BERT, such as XLNet and ERNIE, have demonstrated state-of-the-art performance on numerous benchmark datasets, highlighting the effectiveness of Transformer-based approaches in sentiment analysis [18][19].

Despite these advancements, sentiment analysis applications in educational contexts remain underexplored, particularly for analyzing student feedback on online learning platforms. Previous studies have primarily focused on business and social media domains, with limited datasets available for educational settings. To address this gap, I propose integrating Fuzzy C-Means (FCM) clustering with CNN for sentiment analysis of student feedback. FCM is a clustering method that allows each data point to belong to multiple clusters with varying degrees of membership, making it particularly suitable for handling the ambiguity and subjectivity inherent in sentiment analysis [20].

Recent advancements in artificial intelligence and secure communication have contributed significantly to improving analytical and security frameworks, which are also crucial for enhancing sentiment analysis models in e-learning. Wu et al. introduced a Spectral Convolutional Neural Network model based on Adaptive Fick's Law, demonstrating the effectiveness of deep learning techniques in handling complex, high-dimensional data [21]. Similarly, advancements in secure communication protocols, such as enhanced authentication mechanisms for IoT devices and lightweight key agreement protocols for smart systems, provide foundational insights into the development of robust, scalable, and reliable systems [22][23]. These studies align with the need for integrating innovative machine learning approaches with secure frameworks to address challenges in real-world applications, including e-learning platforms, where data privacy and analytical accuracy are paramount.

3. **Methodology.**

In this study, I propose a hybrid model combining Gated Recurrent Units (GRU) with FCM clustering to analyze the sentiment of comments from online courses. The model architecture is depicted in Figure 1.

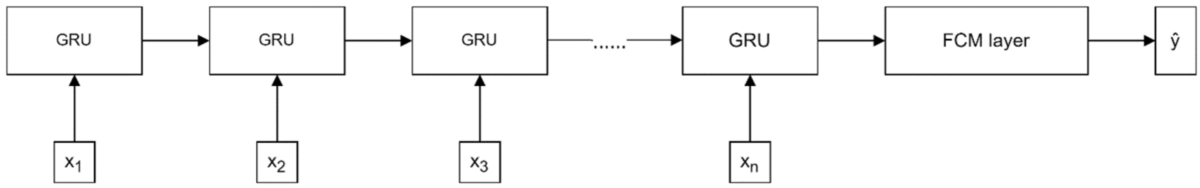

Figure 1. Network structure of GRU-FCM model

The input to the model consists of sequences of comments $\{x_1, x_2, x_3, \ldots, x_n\}$, where each comment is represented as a vector of features. These sequences are fed into a series of GRU layers, which are designed to capture temporal dependencies and contextual information within the comment sequences. The GRU, an advanced variant of RNN, efficiently handles the vanishing gradient problem and is well-suited for sequential data processing.

Each GRU layer processes the input sequence, capturing both forward and backward dependencies, and outputs a hidden state representation that encapsulates the contextual sentiment information. Following the GRU layers, the final hidden state is passed to an FCM layer. The FCM algorithm assigns membership degrees to each data point, allowing for soft clustering where comments can belong to multiple sentiment clusters with varying degrees of certainty. This soft clustering capability of FCM is crucial for sentiment analysis as it can handle the ambiguity and overlap in sentiments more effectively than hard clustering methods. The FCM layer produces a final sentiment prediction $\hat{y}$, which aggregates the fuzzy membership values to determine the overall sentiment polarity of the comments. This hybrid approach leverages the sequential learning strengths of GRU and the clustering flexibility of FCM, resulting in a robust model for sentiment analysis in online course review systems.

3.1. **GRU architecture.** The GRU is a variant of RNN that is particularly effective in managing sequential data, making it suitable for tasks such as sentiment analysis. GRUs are designed to preserve critical information from previous observations in the sequence and capture temporal dependencies [24]. This capability is essential for the analysis of sentiment in text data, as the sentiment conveyed by subsequent words is substantially influenced by the context provided by preceding words.

While LSTM networks and GRU are conceptually similar, GRUs are designed to be more computationally efficient and less susceptible to overfitting, particularly when dealing with large datasets or restricted computational resources. By simplifying the architecture without significantly sacrificing performance, GRUs provide an efficient alternative to traditional RNNs and LSTMs in many NLP tasks, including sentiment analysis. This efficacy renders GRUs especially useful for real-time sentiment analysis applications and the management of extensive online course feedback data. The internal network structure of a GRU unit is depicted in Figure 2.

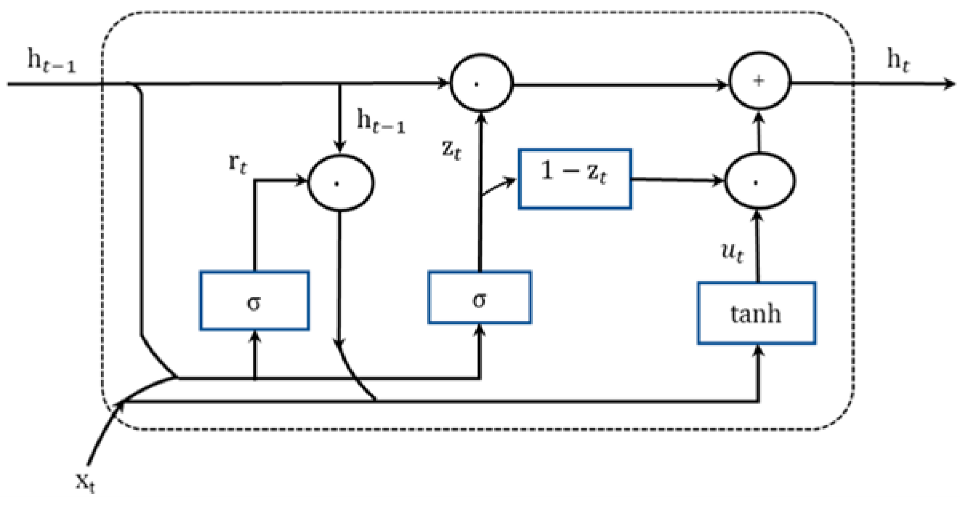

Figure 2. The architecture of GRU unit

The flow of information is regulated by the reset gate r_t and the update gate z_t in the GRU architecture. The input state is represented by x_t, while the previous and current concealed states are represented by h_{t-1} and h_t, respectively. The following equations are used by the GRU to process these states:

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z) \tag{1}$$

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \tag{2}$$

$$\tilde{h}_t = \tanh(W_h x_t + U_h(r_t \odot h_{t-1}) + b_h) \tag{3}$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \tag{4}$$

The update gate z_t is calculated by Equation (1), which determines the amount of the previous hidden state that should be preserved and the amount that should be supplanted by the new information. The input is compressed into the range $[0, 1]$ by the sigmoid function, resulting in a value between 0 and 1.

The reset gate r_t is calculated by Formula (2), which determines the amount of the preceding hidden state that should be discarded and the amount that should be used to calculate. Similar to z_t, it is also a value between 0 and 1.

Equation (3) calculates the candidate hidden state $\tilde{h}_t$, which is a new value that is combined with the preceding hidden state using the update gate. The input is compressed into the range $[-1, 1]$ by the hyperbolic tangent function. Finally, Formula (4) calculates the new hidden state by combining the previous hidden state, h_{t-1}, and the candidate hidden state, $\tilde{h}_t$, in a linear fashion, with the update gate's weight factored in. The symbol $\odot$ denotes element-wise multiplication, while $1 - z_t$ represents the complement of z_t. This guarantees that the new concealed state exclusively incorporates pertinent information from both the previous state and the candidate state.

The GRU unit's architecture, which includes gating mechanisms, enables improved oversight of the transmission of information through the network and has been demonstrated to be exceptional in a variety of tasks. GRU and FCM are combined to create a robust framework for sentiment analysis, which captures the complex temporal dependencies in text data and enables the ambiguous classification of sentiments. The proposed model is highly adaptable and suitable for the analysis of complex sentiment information gathered from online course reviews, as it effectively utilizes the strengths of both GRU and FCM.

3.2. **Fuzzy C-Means.** A clustering method known as FCM separates samples into multiple fuzzy categories, enabling each sample to be a member of multiple clusters with variable degrees of membership. FCM manages the degree of membership for each sample in each category, in contrast to conventional rigid clustering methods. In order to resolve the inherent ambiguity and overlap in sentiment classification, I have integrated FCM with a GRU-based sentiment analysis model in this study. The Fuzzy C-Means algorithm is described as follows:

Each sample is represented as x_i, and each cluster center as v_j, given N samples and C cluster centers. I establish the fuzzy membership matrix U, which U_{ij} denotes the degree of x_i sample membership in the v_j cluster center. The objective of FCM is to reduce the cost function in the following manner:

$$J(U, V) = \sum_{i=1}^{N} \sum_{j=1}^{C} U_{ij}^m \|x_i - v_j\|^2 \tag{5}$$

Here, the fuzziness parameter m regulates the level of fuzziness in the clustering procedure. The weighted sum of squared Euclidean distances between samples and cluster centers is aggregated by this cost function. The algorithm executes the subsequent procedures:

1. **Initialization:** Initialize the fuzzy membership matrix U and the cluster centers V.
2. **Iterative Updates:** Perform iterative updates until a stopping criterion is satisfied.

- **Update** U: Calculate the level of membership of each sample in each cluster center by employing the update formula:

$$U_{ij} = \left(\sum_{k=1}^{C} \left(\frac{\|x_i - v_j\|}{\|x_i - v_k\|} \right)^{\frac{2}{m-1}} \right)^{-1} \tag{6}$$

- **Update** V: Update the cluster centers using the following formula:

$$v_j = \frac{\sum_{i=1}^{N} U_{ij}^m x_i}{\sum_{i=1}^{N} U_{ij}^m} \tag{7}$$

3. **Output:** Return the final fuzzy membership matrix U and the cluster centers V.

The iterative modifications of the fuzzy membership matrix and the cluster centers are the fundamental components of this algorithm. The objective function is minimized through the continuous adjustment of cluster center positions and membership degrees. The fuzzy membership matrix enables us to ascertain the fuzzy membership degrees of each sample in relation to each cluster center, thereby enabling a more nuanced sentiment classification as opposed to rigid categories.

4. Experiment Results.

4.1. Experiment Metric.

To evaluate the effectiveness of the proposed hybrid GRU and Fuzzy C-Means (FCM) model for sentiment analysis, a comprehensive series of experiments were conducted. These experiments covered various aspects, including dataset selection, evaluation metrics, and hyperparameter tuning. The performance of the model was assessed using standard metrics such as true positive (TP), true negative (TN), false positive (FP), and false negative (FN) rates.

The average accuracy was calculated as the ratio of correctly predicted instances to the total number of instances, providing an overall measure of the model's prediction capability. Precision, defined as the ratio of true positive predictions to the sum of true positive and false positive predictions, was used to evaluate the correctness of positive predictions.

Recall, or sensitivity, measured the proportion of true positive instances correctly identified among all actual positive instances. To provide a balanced evaluation of precision and recall, the F1 score was utilized. The F1 score, the harmonic mean of precision and recall, offered a comprehensive metric that combined both aspects, ensuring a robust assessment of the model's performance.

These metrics were chosen because they provide a comprehensive evaluation of the model's predictive performance. Accuracy measures the overall correctness of predictions, while precision focuses on the proportion of true positive predictions among all positive predictions. Recall evaluates the model's ability to identify all relevant instances, and F1 score offers a balanced measure of precision and recall, especially useful in imbalanced datasets. Their combined use ensures a robust assessment of the model's effectiveness in sentiment analysis tasks, where both precision and recall are critical for capturing nuanced feedback from students.

The objective of this study was to evaluate the influence of a new multilayer attention mechanism on a sentiment analysis model. In order to achieve this objective, a sequence

Table 1. Confusion Matrix

Actual / Predict	Positive	Negative
Positive	TP	FP
Negative	FN	TN

of experiments was carried out, which involved the careful selection of suitable datasets, assessment metrics, and hyperparameter setups. To evaluate the performance of the proposed model, the average accuracy and F1 score were calculated using Equations (8) through (11).

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{8}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{9}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{10}$$

$$F_{\text{1score}} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{11}$$

4.2. **Experiment Setup.** In this study, I utilize a novel dataset derived from MOOCs platforms, including 中国大学 MOOC (Chinese University MOOC). This dataset encompasses student feedback on a wide range of online courses, providing a comprehensive overview of student sentiments towards different aspects of E-learning in China. The dataset was collected through an automated web scraping process, using custom-built Python scripts leveraging the BeautifulSoup and Selenium libraries. Feedback comments were extracted from course review sections on Chinese University MOOC [**?**]. These platforms were chosen due to their extensive course offerings and large user bases, ensuring a diverse and abundant dataset. The dataset consists of 12,324 student reviews, with each review including the following fields:

- **Review ID**: A unique identifier for each review.
- **Course ID**: Identifier for the course the review pertains to.
- **Platform**: The E-learning platform from which the review was collected.
- **Review Text**: The textual content of the student's feedback.
- **Rating**: A numerical rating provided by the student, ranging from 1 to 5.

The reviews are categorized into three sentiment classes: positive, neutral, and negative. This classification was performed using a combination of manual annotation and automated sentiment analysis tools, ensuring high accuracy in sentiment labels.

To prepare the data for analysis, several preprocessing steps were undertaken:

- **Text Cleaning**: Removing HTML tags, special characters, and redundant whitespaces.
- **Tokenization**: Splitting the review text into individual tokens (words).
- **Stopword Removal**: Eliminating common stopwords that do not contribute to sentiment (e.g., "和", "的", "是").
- **Lemmatization**: Converting words to their base forms to ensure consistency.
- **Label Encoding**: Transforming sentiment labels into numerical format for use in machine learning models.

The choice of lemmatization over stemming was specifically made to maintain semantic accuracy, especially for Chinese text. In Chinese, stemming may truncate words into forms that lack morphological clarity, leading to a potential loss of context and meaning. Lemmatization ensures that words retain their correct base forms, which is essential for sentiment analysis tasks requiring a nuanced understanding of text. This decision was particularly relevant given the dataset's reliance on accurate sentiment extraction from student feedback.

To illustrate the content and structure of the dataset, here are a few anonymized examples of student reviews:

Example 1

- Review ID: 12345
- Course ID: 67890
- Platform: 中国大学 MOOC
- Review Text: "这门课非常有帮助！老师解释复杂的概念非常清楚。"
- Rating: 5
- Sentiment: Positive

Example 2

- Review ID: 54321
- Course ID: 98765
- Platform: 中国大学 MOOC
- Review Text: "内容不错，但讲解可以更好。有些部分讲得太快了。"
- Rating: 3
- Sentiment: Neutral

Example 3

- Review ID: 11223
- Course ID: 44556
- Platform: 中国大学 MOOC
- Review Text:"我觉得这门课非常无聊，难以跟上。例子也不太相关。"
- Rating: 1
- Sentiment: Negative

The dataset comprised 12,324 comments, with a distribution of sentiments as follows: 4,102 positive, 4,150 neutral, and 4,072 negative comments. The dataset was partitioned into three subsets: a training set, a validation set, and a test set, distributed in proportions of 80%, 10%, and 10% correspondingly. Figure 3 provides comprehensive information about the distribution details of the dataset. The primary aim of this analysis was to evaluate the sentiment conveyed in course comments and detect patterns in sentiment trends among senior students. The primary objective was to improve the standard of instruction and foster active participation from students.

The experimental setup for evaluating the proposed hybrid GRU and FCM model for sentiment analysis was conducted in a high-performance computing environment to ensure efficient processing and accurate results. The system environment included an Intel Xeon Gold 6230R CPU 2.10GHz with 26 cores, 96 GB RAM, an NVIDIA Tesla V100 GPU with 32 GB memory, and 2 TB NVMe SSD for storage. The software configuration comprised Ubuntu 20.04 LTS as the operating system, Python 3.8.10, and the PyTorch 1.8.1 deep learning framework. Essential libraries and dependencies used were NumPy 1.19.5, Matplotlib 3.3.4, Scikit-learn 0.24.1, Pandas 1.2.4, and CUDA Toolkit 11.2 for GPU acceleration. The training parameters for the hybrid GRU and FCM model are detailed in Table 2.

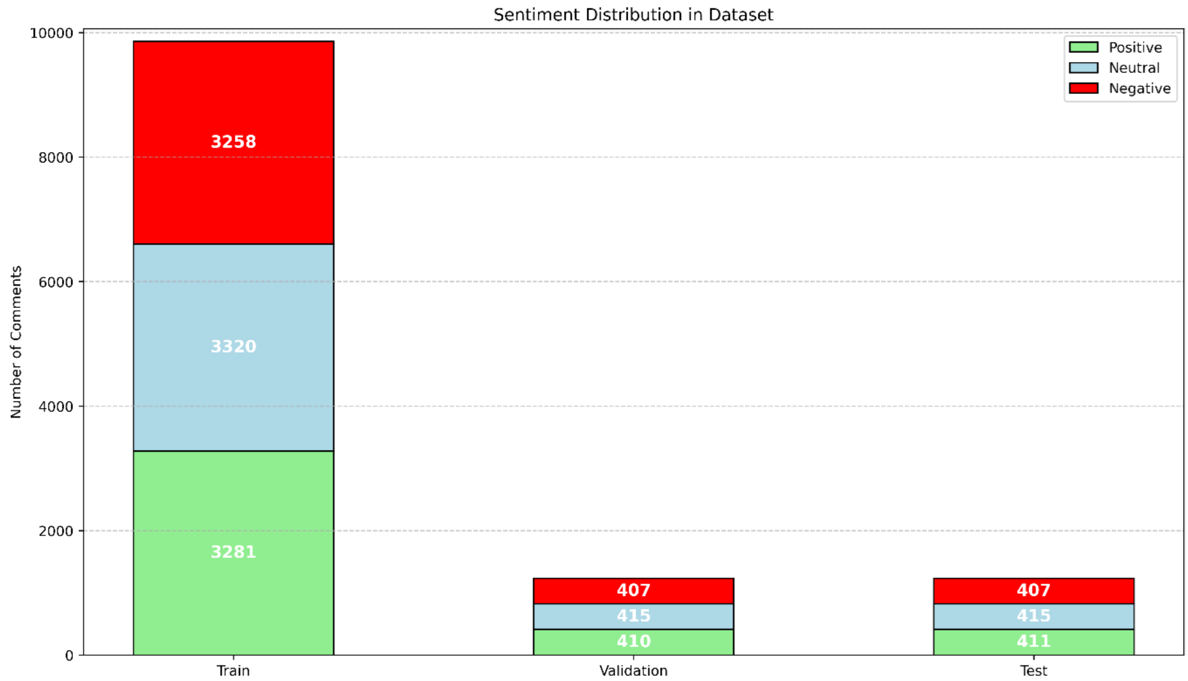

Figure 3. Dataset Description

Table 2. Training parameters

Parameter	Value
Optimizer	Adam
Weight decay	3e-5
Learning rate	0.01
GRU Hidden Layer	256
Number of GRU Layers	2
Dropout	0.5
Number of Clusters (FCM)	3
Fuzziness Parameter (m)	2

4.3. **Experimental Results and Analysis.** The proposed GRU-FCM model was trained over 30 epochs to achieve a reliable level of accuracy. Throughout the training and validation phases, the model exhibited satisfactory loss levels, indicating its effectiveness and stability. The training and validation loss for the GRU-FCM model are illustrated below:

The performance of the proposed model was assessed using key evaluation metrics, namely accuracy and precision. Throughout the training and evaluation phases, the model consistently demonstrated high effectiveness. To validate the efficacy of the proposed model, a comparative analysis was performed against four baseline models, focusing on evaluating the respective outcomes. The proposed model outperformed the baseline models, particularly in terms of accuracy. The comparative accuracy of the five methodologies is illustrated in Figure 5.

Precision is a statistic used to assess the effectiveness of a learning model. It is calculated by dividing the number of true positives (TP) by the total number of predictions made by the model. The proposed model shown improved performance compared to previous models in terms of precision calculation and analysis. Figure 6 depicts the precision comparison of all the examined models.

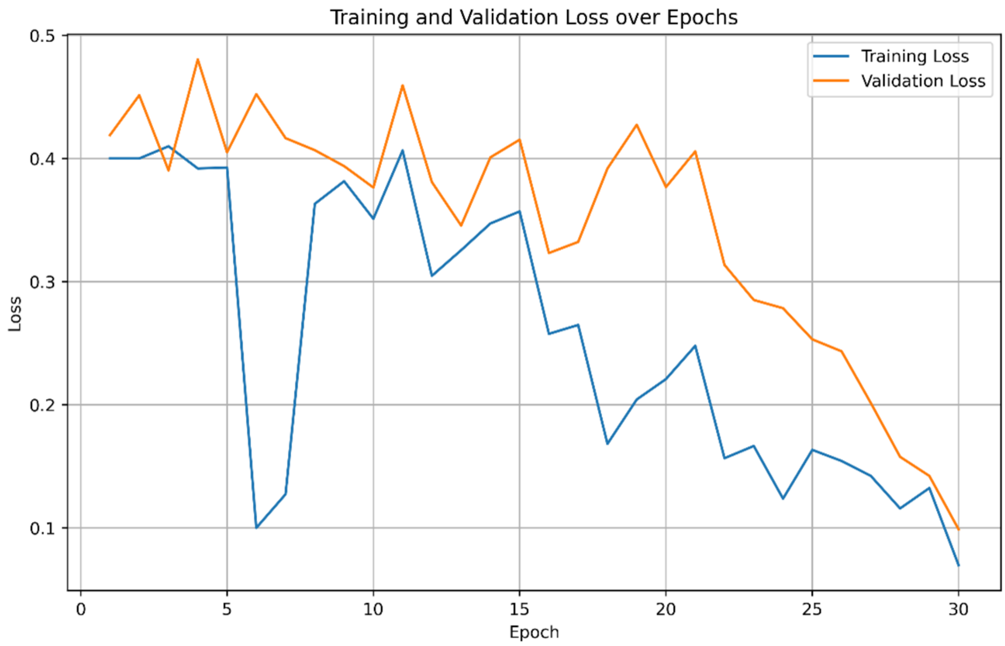

Figure 4. Training and validation loss of GRU-FCM

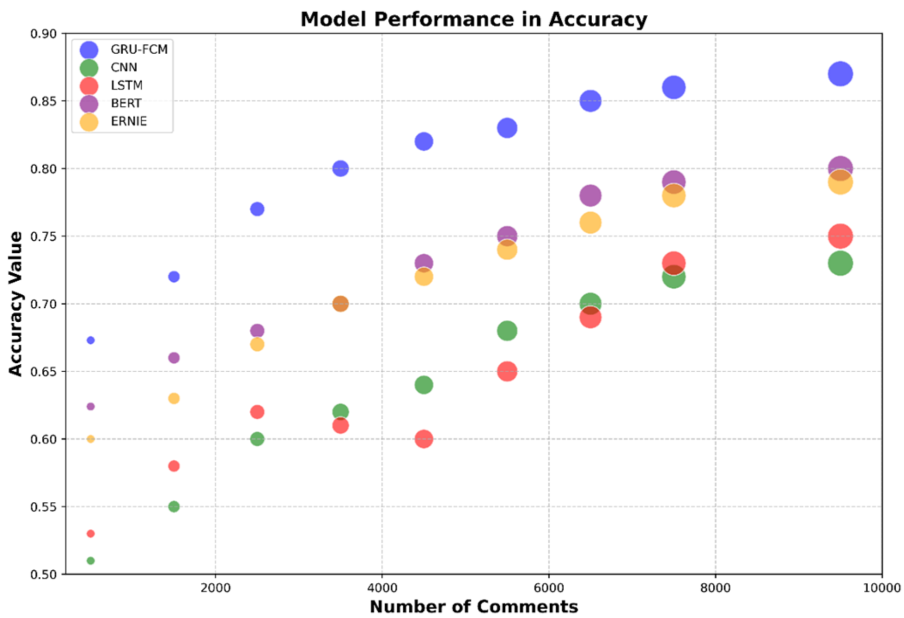

Figure 5. Accuracy of GRU-FCM and four baseline models

F1-score and recall are evaluation measures used to determine the optimum model, together with accuracy and precision. Table 3 and Figure 7 present a comparison of models based on all evaluation metrics.

Table 3 and Figure 7 present a performance comparison of various sentiment analysis models: CNN, LSTM, BERT, ERNIE, and the proposed GRU-FCM model. The GRU-FCM model demonstrates superior performance across all metrics, achieving the highest

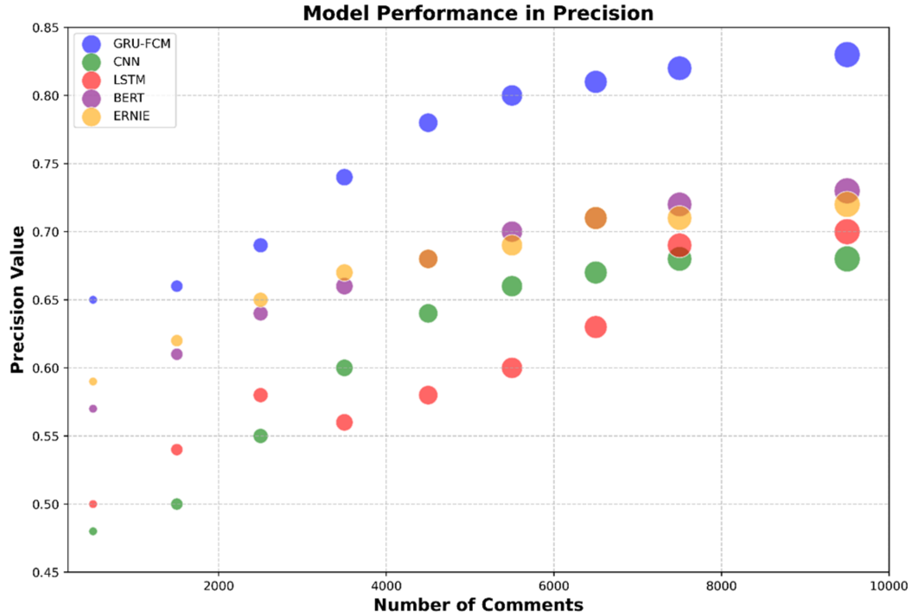

Figure 6. Precision of GRU-FCM and four baseline models

Table 3. Comparative metrics

Methods	Acc	Precision	Recall	F1
CNN	0.73	0.68	0.74	0.72
LSTM	0.75	0.70	0.81	0.74
BERT	0.80	0.73	0.85	0.77
ERNIE	0.79	0.72	0.82	0.75
GRU-FCM	0.87	0.83	0.90	0.86

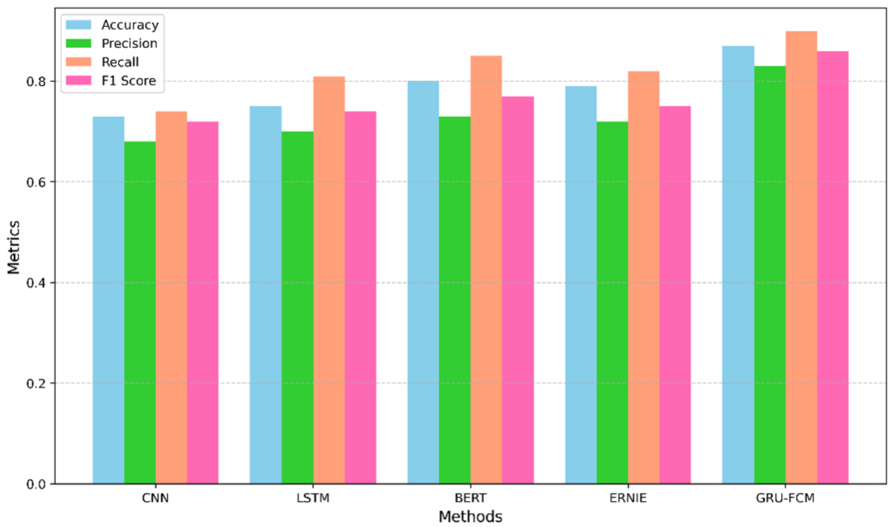

Figure 7. Comparative Metrics by Method

accuracy (0.87), precision (0.83), recall (0.90), and F1 score (0.86). This model outperforms traditional models like CNN (accuracy 0.73, F1 score 0.72) and LSTM (accuracy

0.75, F1 score 0.74), as well as more advanced models like BERT (accuracy 0.80, F1 score 0.77) and ERNIE (accuracy 0.79, F1 score 0.75). The significant improvements in precision and recall indicate that GRU-FCM not only reduces false positives but also captures a larger proportion of positive instances correctly. The high F1 score suggests a balanced and effective handling of both precision and recall. These results underscore the robustness and reliability of the GRU-FCM model for sentiment analysis tasks, particularly in the context of online course reviews, highlighting its ability to accurately predict sentiments and outperform other state-of-the-art models. To further validate the model's effectiveness, statistical significance tests were conducted to compare the GRU-FCM model with baseline models. The results indicated that the improvements in accuracy, precision, recall, and F1 score were statistically significant ($p < 0.05$), confirming that the observed enhancements were unlikely due to random variation. These findings underscore the robustness and reliability of the proposed model in sentiment analysis tasks.

4.4. **Comparing to The Other Related Works.** To conduct a comprehensive empirical assessment and analysis, the outcomes will be compared with contemporary research. Subsequently, an evaluation will be performed to assess the efficacy of five recent models: our proposed GRU-FCM, SMA (Semantic Analysis model) [26], FeD (Feeling Distinguisher system) [27], LSTM-Attention [28], ASDA (Aspect-Based Sentiment Analysis) [29] (2022), and TextCNN-CAT-FEW (Channel Attention TextCNN with Feature Word Extraction for Chinese Sentiment Analysis) [30] (2023). Table 4 presents this comparison.

Table 4. Comparative experimental findings with five recent models in the literature

Methods	**Accuracy**	**Precision**	**Recall**	**F1**
SMA [26]	-	0.76	0.94	0.84
FeD [27]	-	0.71	0.87	0.78
LSTM-Attention [28]	-	0.69	0.82	0.75
ASDA [29]	0.81	-	-	-
TextCNN-CAT-FEW [30]	0.83	0.81	0.80	0.80
GRU-FCM	0.87	0.83	0.90	0.86

Table 4 presents a comparative analysis of the proposed GRU-FCM model against five recent sentiment analysis models: SMA [26], FeD [27], LSTM-Attention [28], ASDA [29], and TextCNN-CAT-FEW [30]. The GRU-FCM model outperforms all others, achieving the highest accuracy (0.87), precision (0.83), recall (0.90), and F1 score (0.86). This indicates its superior capability in reducing false positives and correctly identifying positive instances. In contrast, the other models show lower metrics, with ASDA and TextCNN-CAT-FEW having accuracies of 0.81 and 0.83, respectively, and SMA, FeD, and LSTM-Attention having lower precision and F1 scores. The results underscore the effectiveness and robustness of the GRU-FCM model for sentiment analysis, making it highly suitable for analyzing sentiments in online course reviews.

4.5. **Discussion.** In response to increasing concerns about data privacy and potential bias in sentiment analysis, we have implemented measures to address these issues throughout the study. To protect student privacy, all data were anonymized, and access to sensitive information was strictly controlled. To ensure fairness, we balanced the dataset to represent diverse student populations, preventing the model from unintentionally favoring or disadvantaging any specific group. The preprocessing steps were also designed carefully to remove any unintended discriminatory language that could introduce bias into the analysis. Furthermore, the model was tested across various demographic subsets to verify that its predictions remained consistent and unbiased. By incorporating these measures, we aim to provide not only a robust sentiment analysis model but also one that upholds fairness, inclusivity, and ethical standards in processing student feedback.

Future research could focus on incorporating advanced neural architectures, such as Transformer-based models or graph neural networks, to further improve the model's capacity for understanding complex patterns in sentiment analysis. Additionally, applying the model to multilingual datasets or adapting it to different cultural and educational contexts could provide valuable insights into its generalizability and scalability. Another promising direction would be to explore its use in real-time analysis of dynamic interactions, such as live classroom discussions or synchronous online learning sessions. These advancements, combined with efforts to gather more diverse and representative datasets, could significantly enhance the model's applicability and effectiveness in addressing the evolving challenges of e-learning sentiment analysis.

5. **Conclusion.** In this study, I proposed a hybrid GRU-FCM model for sentiment analysis of student feedback on e-learning platforms, achieving superior performance compared to traditional and state-of-the-art models such as CNN, LSTM, BERT, and ERNIE. The GRU-FCM model demonstrated the highest accuracy (0.87), precision (0.83), recall (0.90), and F1 score (0.86), underscoring its robustness and reliability. This model effectively captures temporal dependencies and manages sentiment ambiguity, making it highly effective for practical applications in educational settings to improve e-learning experiences. The comprehensive empirical assessment and comparison with recent models validate the GRU-FCM's efficacy in sentiment analysis. Future work can explore the integration of additional neural architectures and advanced clustering techniques, as well as expanding the dataset to include more diverse feedback from various e-learning platforms, to further enhance the model's generalizability and performance.

REFERENCES

[1] R. Hassan, F. Qamar, M. K. Hasan, A. H. M. Aman, and A. S. Ahmed, "Internet of Things and Its Applications: A Comprehensive Survey," *Symmetry*, vol. 12, no. 10, Art. no. 10, Oct. 2020, doi: 10.3390/sym12101674.

[2] "The COVID-19 pandemic has changed education forever. This is how," *World Economic Forum*, Apr. 2020, . Available: `https://www.weforum.org/agenda/2020/04/coronavirus-education-global-covid19-online-digital-learning/`

[3] M. Radović-Marković, "Advantages And Disadvantages Of E-Learning In Comparison To Traditional Forms Of Learning," *Annals of the University of Petrosani, Economics*, vol. 10, no. 2, Art. no. 2, 2010.

[4] W. He, G. Xu, and S. Kruck, "Online IS Education for the 21st Century," *Journal of Information Systems Education*, vol. 25, no. 2, Art. no. 2, Jan. 2014.

[5] R. Baragash and H. Aldowah, "Sentiment analysis in higher education: a systematic mapping review," *Journal of Physics: Conference Series*, vol. 1860, no. 1, Art. no. 1, Mar. 2021, doi: 10.1088/1742-6596/1860/1/012002.

[6] B. Pang and L. Lee, "A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts," in *Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics (ACL-04)*, Barcelona, Spain, Jul. 2004, pp. 271–278. doi: 10.3115/1218955.1218990.

[7] Z. Khanam, "Sentiment Analysis of user reviews in an Online Learning Environment: Analyzing the Methods and Future Prospects," *European Journal of Education and Pedagogy*, vol. 4, pp. 209–217, Apr. 2023, doi: 10.24018/ejedu.2023.4.2.531.

[8] H. J. Alantari, I. S. Currim, Y. Deng, and S. Singh, "An empirical comparison of machine learning methods for text-based sentiment analysis of online consumer reviews," *International Journal of Research in Marketing*, vol. 39, no. 1, pp. 1–19, 2022.

[9] F. Dalipi, K. Zdravkova, and F. Ahlgren, "Sentiment Analysis of Students' Feedback in MOOCs: A Systematic Literature Review," *Frontiers in Artificial Intelligence*, vol. 4, p. 728708, Sep. 2021, doi: 10.3389/frai.2021.728708.

[10] T. Luong, H. Pham, and C. D. Manning, "Effective Approaches to Attention-based Neural Machine Translation," in *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, Lisbon, Portugal, Sep. 2015, pp. 1412–1421. doi: 10.18653/v1/D15-1166.

[11] Z. Yang et al., "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, California, Jun. 2016, pp. 1480–1489. doi: 10.18653/v1/N16-1174.

[12] S. Kumar et al., "Sentiment Analysis on Online Reviews using Machine Learning and NLTK," in *2022 6th International Conference on Trends in Electronics and Informatics (ICOEI)*, Apr. 2022, pp. 1183–1189. doi: 10.1109/ICOEI53556.2022.9776850.

[13] E. Cambria et al., "Towards a Chinese Common and Common Sense Knowledge Base for Sentiment Analysis," in *Advanced Research in Applied Artificial Intelligence*, Berlin, Heidelberg: Springer, 2012, pp. 437–446. doi: 10.1007/978-3-642-31087-4_46.

[14] Y. Kim, "Convolutional Neural Networks for Sentence Classification," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Doha, Qatar, Oct. 2014, pp. 1746–1751. doi: 10.3115/v1/D14-1181.

[15] Y. Zhou et al., "Text Sentiment Analysis Based on a New Hybrid Network Model," *Computational Intelligence and Neuroscience*, vol. 2022, p. 6774320, Dec. 2022, doi: 10.1155/2022/6774320.

[16] A. Chanaa and N. El Faddouli, "Sentiment Analysis on Massive Open Online Courses (MOOCs)," *IGI Global*, Jan. 2023. `https://www.igi-global.com/article/sentiment-analysis-on-massive-open-online-courses-moocs/310004`

[17] J. Devlin et al., "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics*, Minneapolis, Minnesota, Jun. 2019, pp. 4171–4186. doi: 10.18653/v1/N19-1423.

[18] Z. Yang et al., "XLNet: Generalized Autoregressive Pretraining for Language Understanding," in *Proceedings of the 33rd International Conference on Neural Information Processing Systems*, Vancouver BC Canada, Dec. 2019. `https://papers.nips.cc/paper_files/paper/2019/hash/dc6a7e655d7e5840e66733e9ee67cc69-Abstract.html`

[19] Y. Sun et al., "ERNIE 2.0: A Continual Pre-Training Framework for Language Understanding," *AAAI Conference on Artificial Intelligence*, vol. 34, Apr. 2020, doi: 10.1609/aaai.v34i05.6428.

[20] J. C. Bezdek et al., "FCM: The fuzzy c-means clustering algorithm," *Computers & Geosciences*, vol. 10, no. 2, pp. 191–203, Jan. 1984, doi: 10.1016/0098-3004(84)90020-7.

[21] T.-Y. Wu, H. Li, S. Kumari, and C.-M. Chen, "A Spectral Convolutional Neural Network Model Based on Adaptive Fick's Law for Hyperspectral Image Classification," *Computers, Materials & Continua*, vol. 79, no. 1, pp. 19–46, 2024.

[22] T.-Y. Wu, F. Kong, Q. Meng, S. Kumari, and C.-M. Chen, "Rotating Behind Security: An enhanced authentication protocol for IoT-enabled devices in distributed cloud computing architecture," *EURASIP Journal on Wireless Communications and Networking*, vol. 2023, p. 36, 2023.

[23] T.-Y. Wu, L. Wang, and C.-M. Chen, "Enhancing the Security: A Lightweight Authentication and Key Agreement Protocol for Smart Medical Services in the IoHT," *Mathematics*, vol. 11, no. 17, p. 3701, 2023.

[24] G. Shen, Q. Tan, H. Zhang, P. Zeng, and J. Xu, "Deep Learning with Gated Recurrent Unit Networks for Financial Sequence Predictions," *Procedia Computer Science*, vol. 131, pp. 895–903, Jan. 2018, doi: 10.1016/j.procs.2018.04.298.

[25] P. Liu, X. Qiu, and X. Huang, "Recurrent Neural Network for Text Classification with Multi-Task Learning," *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, NY USA, May 2016.

[26] L. Wang, G. Hu, and T. Zhou, "Semantic Analysis of Learners' Emotional Tendencies on Online MOOC Education," *Sustainability*, vol. 10, p. 1921, Jun. 2018, doi: 10.3390/su10061921.

[27] C.-H. Tai, Z.-H. Tan, and Y.-S. Chang, "Systematical Approach for Detecting the Intention and Intensity of Feelings on Social Network," *IEEE Journal of Biomedical and Health Informatics*, vol. 20, no. 4, Jul. 2016, doi: 10.1109/JBHI.2016.2535721.

[28] K. S. Tai, R. Socher, and C. D. Manning, "Improved Semantic Representations From Tree-Structured Long Short-Term Memory Networks," 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, Jul. 2015, pp. 1556–1566. doi: 10.3115/v1/P15-1150.

[29] J. Melba Rosalind and S. Suguna, "Predicting Students' Satisfaction Towards Online Courses Using Aspect-Based Sentiment Analysis," in *Computer, Communication, and Signal Processing*, Cham: Springer International Publishing, 2022, pp. 20–35. doi: 10.1007/978-3-031-11633-9_3.

[30] J. Liu, Z. Yan, S. Chen, X. Sun, and B. Luo, "Channel Attention TextCNN with Feature Word Extraction for Chinese Sentiment Analysis," *ACM Transactions on Asian and Low-Resource Language Information Processing*, vol. 22, no. 4, p. 100:1–100:23, Mar. 2023, doi: 10.1145/3571716.