Journal of Network Intelligence ©2024 ISSN 2414-8105 (Online)
Taiwan Ubiquitous Information Volume **9**, Number **3**, August **2024**

Sentiment Analysis of Student Comment Text on Online Education Platform by Fusing CNN and Long And Short Term Memory Network

Zhongmin Li

School of Foreign Languages
Hunan University of Arts and Science, Changde 415000, China
lzm1988jimmy@huas.edu.cn

Jinxia Lei*

School of Education Science
Zhoukou Normal University, Zhoukou 466600, China
ayjinxialei@hotmail.com

*Corresponding author: Jinxia Lei

Received October 29, 2023, revised February 15, 2024, accepted May 4, 2024.

Abstract. *In the rapidly evolving landscape of online education, effective teacher-student engagement is crucial. This study addresses the challenge of autonomously analyzing student feedback in online learning environments. We introduce the Multi-Attention Mechanism (MAM), a novel sentiment analysis model that combines Convolutional Neural Networks (CNN) and Long and Short Term Memory Network (LSTM). MAM excels in identifying the affective polarity of student comments by using a gate mechanism for precise affective word detection and a hierarchical attention mechanism for handling diverse emotional expressions. It also integrates original data into the attention mechanism to minimize information loss. Rigorous testing on datasets from MOOC and Tencent Classroom in Chinese universities shows that MAM significantly outperforms traditional models in Accuracy and Precision at 0.85 and 0.81. This research advances educational technology by improving the analysis of student feedback, ultimately enhancing the quality of online education.*
Keywords: Information Systems and Applications, Emerging Technologies and Applications, Artificial Intelligence, Online platforms, Database Management, Sentiment analysis, Multi-Attention Mechanism.

---

1. **Introduction.** In recent years, the rapid advancement of networking technologies, such as the Internet, Internet of Things (IoT), and big data, has significantly influenced various aspects of our lives, including work, education, and lifestyle [1]. These technologies have paved the way for the emergence of networking platforms that cater to e-commerce, social communication, and education, thereby expanding the boundaries of our daily experiences. The COVID-19 pandemic's worldwide spread has necessitat-ed numerous countries to shift from conventional in-person teaching to virtual learn-ing. Online education offers several advantages over conventional classroom-based learning, including reduced costs, increased flexibility, and diminished geographical limitations [2][3][4]. Figure 1 illustrates an exemplar of online education in the form of a MOOC, which has gained substantial popularity due to its accessibility, affordability, and scalability. MOOCs are

designed to facilitate unlimited participation from learners worldwide, granting them access to high quality educational content and enabling interaction with instructors and peers within a vir-tual environment. MOOCs harness a wide range of instructional technologies and methodologies, encompassing video lectures, interactive quizzes, discussion forums, and collaborative assessments, to enhance learning outcomes and foster active engagement among learners. These pedagogical approaches are strategically employed to maximize the efficacy of online education, promoting knowledge acquisition and meaningful interactions in a dynamic virtual setting. In order to assess the caliber of instruction

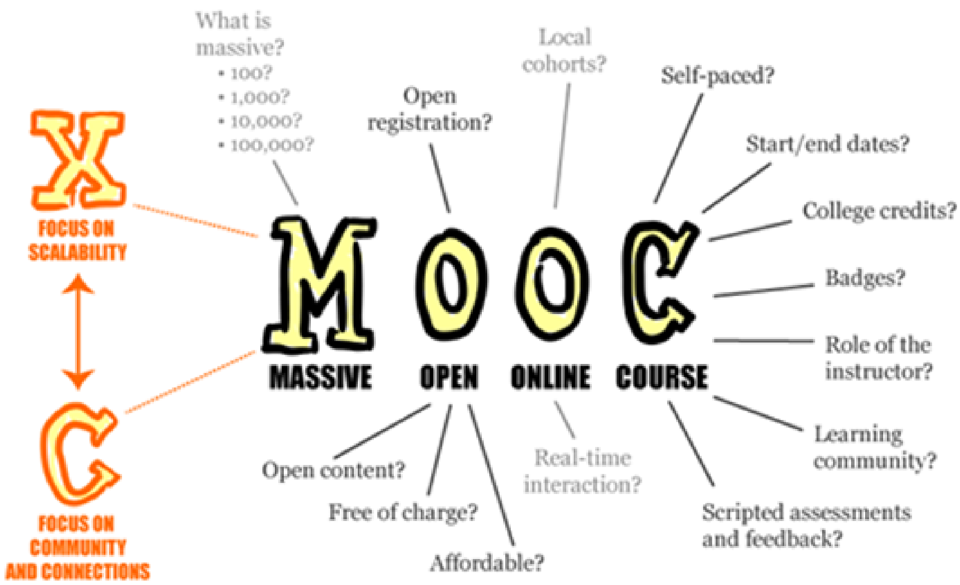

Figure 1. Massive open online course (MOOC describe)

and augment the level of engagement between educators and learners, prominent educational platforms such as China University's MOOC and Tencent Classroom have incorporated the functionality of bullet-pointed chat. Within the online networking environment, emotive comments reflect the perspectives of individuals, and effectively harnessing these viewpoints assumes critical significance in augmenting service quality. Bulleted conversations assume a pivotal role within the pedagogical process, as they enable teachers to identify areas of student weakness through their feedback, facilitating subsequent adjustments in knowledge scores, teaching plans, learning objectives, and faculty composition [5][6].

Sentiment analysis, crucial for assessing textual data's emotional tone, is widely used in sectors like business, healthcare, politics, and education, where it aids in understanding diverse reactions ranging from patient feedback to political sentiments and market trends [20][21][22]. Despite its broad application, challenges persist, including ethical concerns around privacy and bias, and the inherent subjectivity in interpreting emotional content. This necessitates ongoing development of robust, ethically minded sentiment analysis methods that accurately capture emotional nuances while safeguarding marginalized groups [23][24].

Academic researchers have proposed three primary approaches to conduct sentiment analysis, namely: dictionary and rule-based sentiment analysis, traditional machine learning-based sentiment analysis, and deep learning-based sentiment analysis [7]. While sentiment analysis has found widespread usage in the business domain since Pang et al.'s seminal study on sentiment analysis of movie reviews, its application within the realm of online course reviews remains relatively scarce, and publicly available datasets pertaining to this subject remain exceedingly limited [8][9][10]. Researchers have employed diverse approaches to compute sentiment scores in the context of online courses. Notably, the second method, anticipated to capture sentiment by employing artificially constructed features and employing classifiers like Naive Bayes, maximum entropy, and SVM, has been applied in sentiment analysis [11][12]. Nonetheless, the accuracy of this method hinges upon feature construction and classifier selection, with the quality of feature selection exerting a substantial influence on the fidelity of the testing outcomes. Deep learning methodologies, including RNN, LSTM, and CNN, have been widely employed to address the limitations inherent in traditional machine learning approaches.

Traditional neural network architectures, such as LSTM and attention mechanisms, have also been utilized to enhance model accuracy. Despite the limitations of sequence-based neural networks such as LSTM in handling long sequences and computational memory, attention mechanisms provide the ability to capture output dependencies without the need for explicit consideration of intertextual distances. Numerous studies have integrated classical neural networks with attention mechanisms, resulting in improved model performance [13][14][15]. Sentiment analysis assumes a critical role within the realm of education due to its potential to bolster teaching quality, mitigate dropout rates, and foster sustainable development. The utilization of a hierarchical attention mechanism coupled with a portal mechanism, as proposed in this study, exhibits the capacity to enhance sentiment analysis accuracy when appraising online course evaluations [16]. Additionally, feature engineering encounters challenges concerning large scale workloads and feature sparsity, alongside pervasive issues pertaining to domain adaptability [17]. Consequently, the precision of experimental outcomes is substantially influenced by the quality of feature selection, necessitating further exploration of innovative approaches to tackle these challenges [18][19].

This research proposes a novel Multi-Attention Mechanism (MAM) tailored for sentiment analysis of online course reviews to address the limitations of prior techniques. MAM combines a CNN to extract local features and an LSTM to capture global, long-range information. Existing sentiment analysis methods, although diverse and evolving, often face challenges in accurately capturing the nuances of human emotions, particularly in the context of online education. To assess the quality of instruction and engagement in online courses, educational platforms utilize features like bullet-pointed chat, where emotive comments are critical for enhancing service quality [5][6]. While sentiment analysis has been applied across various domains, including healthcare, politics, and finance [20][21][22], it is challenged by issues of privacy, bias, fairness, and the inherent subjectivity in classifying affective valence [23][24]. Therefore, our research introduces MAM, which uniquely combines CNN and LSTM with a hierarchical attention architecture, to overcome these challenges in analyzing sentiments in online course reviews.The contributions of MAM are three-fold:

1. A localized-global learning framework integrating CNN and LSTM to extract multi-granularity features capturing local and long-range semantic contexts.
2. A gating mechanism to selectively filter emotionally salient information to improve sentiment classification accuracy.
3. Hierarchical attention architecture to preserve original information and handle nuanced linguistic expressions of sentiment.

Experiments conducted on Chinese MOOC datasets demonstrate significant improvements in accuracy and precision compared to baseline methods. This work represents a valuable advancement in sentiment analysis technology to enhance online education platforms.

The remainder of this paper is organized as follows. Section 2 reviews related work. Section 3 presents the methodology. Section 4 describes the experimental setup and analyzes results. Section 5 provides concluding remarks.

2. **Related Work.** Sentiment analysis, a widely explored area in natural language processing (NLP), focuses on categorizing textual data into positive, negative, or neutral sentiments. The advent of deep learning has witnessed the emergence of various sentiment analysis models, including graph based and sequence based approaches. This paper aims to provide a comprehensive overview of recent advancements in sentiment analysis modeling and discuss the merits and drawbacks associated with different methodologies.

Graph-based models employ graph convolutional networks (GCNs) to discern sentiments in textual data. One of the pioneering graph-based models, TextGCN introduced by Yao et al. [25], constructs two graphs to capture the associations between words and documents. The Pointwise Mutual Information (PMI) graph establishes word relationships, while the TFIDF graph portrays document-word associations. Notably, this model exhibits high accuracy when evaluated on multiple standardized datasets. Sub-sequently, numerous graph-based models have been proposed, such as HeteGCN proposed by Ragesh et al. [26], which integrates word prediction features with TextGCN. Another noteworthy model, HyperGAT developed by Yadati et al. [28], transforms textual information into a hypergraph, leading to enhanced classification accuracy. Additionally, TensorGCN devised by Liu et al. [29] employs multiple graphs to capture semantic, syntactic, and contextual information, thereby improving the performance of text classification tasks.

Recent studies in the realm of IoT, security, and data mining have laid a foundational understanding that enriches our approach to sentiment analysis in online education platforms. Wu et. al [30] have significantly contributed to this landscape by enhancing security protocols in Smart Medical Services within the Internet of Healthcare Things (IoHT) through lightweight authentication and key agreement protocols, emphasizing the importance of secure data transmission. This emphasis on security is further echoed in Wu et al.'s exploration of secure smart-home IoT access control, which utilizes a home registration approach to bolster IoT security [31]. Moreover, the advancement in IoT-enabled device authentication in distributed cloud computing architectures by Wu, Kong, Meng, Kumari, and Chen underscores the criticality of robust authentication mechanisms in our increasingly connected world [32]. Complementing these studies, Chen et. al delve into the realm of data mining, specifically in high utility periodic frequent pattern mining in multiple sequences, providing insights into sophisticated data handling techniques [33]. Finally, the exploration by Chen, Gong, and Wu into the impact of technical and leading indicators on stock trends in the IoT context presents a nuanced understanding of IoT's influence beyond conventional domains [34]. These collective works not only fortify the security and data integrity aspects of IoT and cloud systems but also provide a broader perspective on the complex interplay between technology and data analysis, which is crucial for our sentiment analysis methodology in online education platforms.

While graph-based models have demonstrated remarkable performance in the field of sentiment analysis, recent studies have suggested that sequence-based models, particularly those based on the Transformer architecture, achieve superior outcomes [27]. Transformer-based models employ self-attention mechanisms to represent input text as a sequence of vectors. Among these models, the Bidirectional Encoder Representations from Transformers (BERT) model, proposed by Devlin et al. [35], has emerged as the most successful one.

BERT is a pre-trained model that can be fine-tuned for various downstream tasks, including sentiment analysis. Moreover, several modified versions of BERT, such as XLNet [36] and ERNIE [37], have exhibited state-of-the-art performance on numerous benchmark datasets.

Sequence-based models have also been investigated for sentiment analysis purposes. Kim [38] leveraged a CNN to classify sentiments within text. By representing individual words, the text was fed into the CNN network to ascertain its sentiment polarity. Liu et al. [39] employed recurrent neural networks (RNNs) for affective classification. Notably, Wang et al. [40] observed that LSTM outperformed traditional RNNs in sentiment analysis of tweets. Similarly, Barnes et al. [41] utilized a two layer LSTM architecture to classify sentiment in tweets, while Yang et al. [15] applied attention mechanisms to aggregate information from sentences during emotion analysis.

In critically analyzing these developments, our proposed MAM model distinguishes itself by integrating the strengths of both graph-based and sequence-based approaches. MAM leverages the local feature extraction capabilities of CNNs, as seen in sequence-based models, while also incorporating the global contextual understanding characteristic of graph-based models. This unique amalgamation addresses the limitations of existing models, particularly in handling complex sentiment expressions in online course reviews. The integration of a hierarchical attention architecture in MAM further advances the field by providing a more nuanced understanding of sentiment, capturing subtleties often missed by previous models. Therefore, our approach not only builds upon the existing literature but also introduces novel elements to elevate sentiment analysis, particularly in the educational domain.

3. **Methodology.**

When analyzing online course reviews for sentiment, a segment of length n is designated for examination. Furthermore, the sentiment polarity conveyed through bullet screen comments is evaluated during the review analysis process. After obtaining the comment sentence S, each word must be vectorized. In this study, the initialization of each word vector in sentence S is performed randomly, denoted by $S_v = [v^1, v^2, \ldots, v^n] \in \mathbb{R}^{n\times d_w}$ with the size of the word vector dimension d_w.

3.1. **Mutilevel Attention Mechanism (MAM) architecture.**

The extraction of local sentiment features from reviews involves a multi-stage approach using advanced deep-learning techniques. First, a hybrid Convolutional Neural Network (CNN) is employed to capture the local features of sentences, represented by the vector $H_C = [h_c^1, h_c^2, \ldots, h_c^m] \in \mathbb{R}^{m\times d_h}$. This approach is illustrated in Figure 2.

Next, the Long Short-Term Memory (LSTM) model is used to extract the hidden information of the comment text. This enables the capture of temporal dependencies within the re-view data, which can be important for sentiment analysis. Finally, the context hidden state $H_L = [h_l^1, h_l^2, \ldots, h_l^n] \in \mathbb{R}^{n\times d_h}$ is obtained by applying LSTM to the input sequence.

A novel gate mechanism is proposed to screen the important sentiment information from H_C and H_L. Specifically, an average pooling operation is performed on the hidden state of the sentence H_L extracted by LSTM and the local emotion information H_C. This generates vector representations, $h_g^L \in \mathbb{R}^{d_h}$ and $h_g^C \in \mathbb{R}^{d_c}$, respectively. These representations are combined using the gate mechanism to filter the most relevant sentiment information selectively. This approach represents a significant improvement over previous sentiment analysis methods by enabling the extraction of more nuanced sentiment

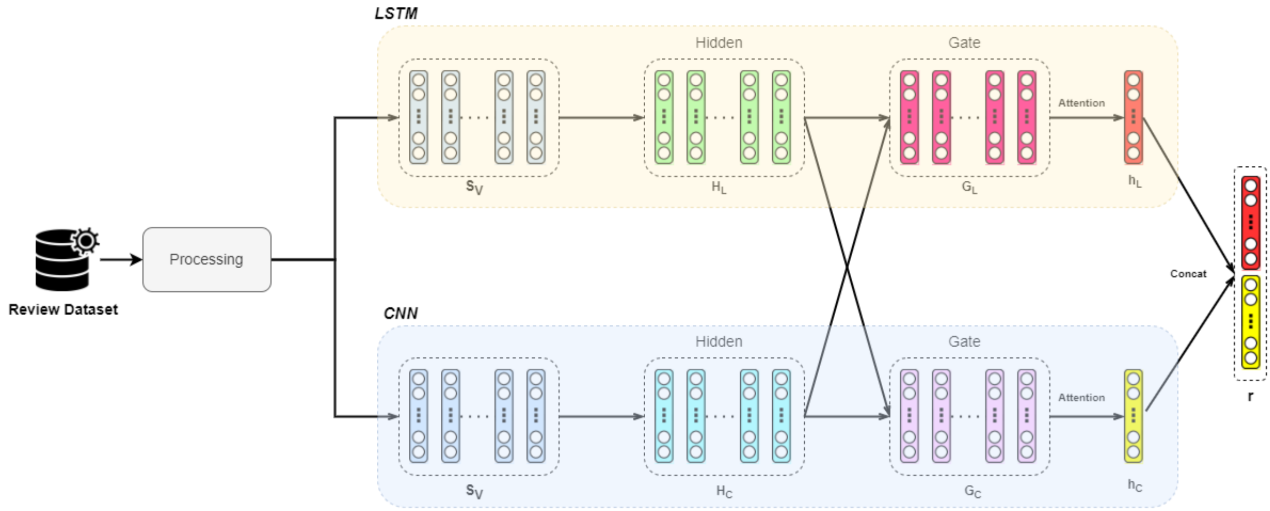

Figure 2. The architecture of MAM.

features and improving sentiment classification accuracy.

$$h_g^L = \sum_{i=1}^{n} \frac{H_L}{n} \tag{1}$$

$$h_C^L = \sum_{i=1}^{m} \frac{H_C}{m} \tag{2}$$

Subsequent to extracting local emotion information from the CNN, a gate mechanism is utilized to filter the information using the vector representation h_g^L. This gating mechanism allows for the selective processing of relevant sentiment features, improving the accuracy of the analysis. The specific calculations involved in this process are detailed in Formulas (3)–(5), providing a clear and concise approach to gating local sentiment information. This technique represents a novel approach to sentiment analysis, enabling more precise classification of emotions and facilitating more informed decision-making in a variety of contexts.

$$T_C = relu\left(H_C W_C + W_g h_g^L \times w_g\right) \tag{3}$$

$$E^i = \tanh\left(H_C^i W_E\right) \tag{4}$$

$$G_C^i = E^i T_C \tag{5}$$

The gate mechanism used to filter local emotion information involves a series of calculations utilizing various parameter vectors and activation functions. Specifically, the gate mechanism involves the use of $W_C \in \mathbb{R}^{d_c \times d_c}, W_g \in \mathbb{R}^{m \times d_h}, W_g \in \mathbb{R}^{d_C}$ as parameter vectors, and relu and tanh as activation functions. In addition, global features $W_E \in \mathbb{R}^{d_c \times m}$ and $G_C^i \in \mathbb{R}^{d_c}$ are utilized in the calculation process. Once the gate mechanism has been selectively applied to the local emotion information, a selective representation of the gate mechanism denoted as $G_C \in \mathbb{R}^{m \times d_c}$, can be obtained by expressing the chosen i vector of H_c through the mechanism. Similarly, the context hidden state $G_L \in \mathbb{R}^{n \times d_h}$ can be obtained through the same approach. Once G_C has been obtained, information is aggregated through an attention mechanism, as demonstrated in Formulas (6) and (7). This approach represents a novel and effective method for sentiment analysis, allowing for more precise identification and classification of emotions in various contexts.

$$\alpha = \text{softmax}\left(G_C w_g + H_C w_C\right) \tag{6}$$

$$h_c = \sum_{i=1}^{m} a^i G_c^i \tag{7}$$

The gate mechanism utilized in this study involves several parameter vectors, including $w_\alpha \in \mathbb{R}^{d_c}$ and $w_C \in \mathbb{R}^{d_c}$, and H_C, representing the feature information extracted by the original CNN. After obtaining the selection information G from the gate mechanism, the original information H_C is further integrated into the process when the attention mechanism coefficients are weighted, helping to prevent the loss of valuable information. The resulting vector representation, $h_C \in \mathbb{R}^{d_h}$, is then weighted using Formula (7).

However, the single-layer attention mechanism has limitations in identifying the most important words in sentiment analysis. To address this issue, we propose the use of a multi-layer attention mechanism, which can more effectively highlight the importance of different words in sentences. In this approach, the first layer of the attention mechanism, represented by h_c, to weight the importance of different words. Subsequently, the information contained within G_L is weighted using h_c, as shown in Formulas (8)–(10). By incorporating a multi-layer attention mechanism, this approach represents a significant improvement over previous sentiment analysis techniques, enabling the more precise identification and classification of important sentiment words in various contexts. These findings may have important implications for the development of more effective sentiment analysis methods in the future.

$$\gamma\left(h_L^i, h_c\right) = \tanh\left(h_L^i W_L h_c^T\right) \tag{8}$$

$$\beta_i = \frac{e^{\gamma\left(h_t^i, h_c\right)}}{\sum_{j=1}^{n} e^{\gamma\left(h_t^j, h_c\right)}} \tag{9}$$

$$h_t = \sum_{i=1}^{n} \beta_i h_t^i \tag{10}$$

The vector representation $h_i^L \in \mathbb{R}^{d_h}$ represents the i-th vector of G_L, and is subjected to the activation function tanh during the calculation process. This calculation involves the use of the parameter matrix $W_L \in \mathbb{R}^{d_h \times d_C}$ and the transpose of the h_c vector, h_c^T. The attention mechanism coefficient $\gamma(h_i^L, h^c)$ is then obtained using Formula (8), providing a measure of the relationship between h_i^L and h_c.

To ensure that the attention mechanism coefficient is normalized, β_i is calculated as the normalized attention mechanism coefficient between h_i^L and h_c. The resulting vector representation, $h_L \in \mathbb{R}^{d_h}$, is obtained through a weighted summation using Formula (10). By splicing the vector representations h_c and h_L, the final representation of a sentence containing sentiment information can be obtained, denoted as $r \in \mathbb{R}^{d_c+d_h}$. Finally, the sentiment polarity of the sentence can be determined through the use of the softmax classifier, as shown in Formulas (11) and (12). This approach represents a significant improvement over previous sentiment analysis methods, enabling more accurate identification and classification of sentiment information in a variety of contexts. These findings may have important implications for the development of more effective sentiment analysis techniques in the future.

$$x = \tanh\left(W_r r + b_r\right) \tag{11}$$

$$y_i = \frac{e^{x_i}}{\sum_{j=1}^{C} e^{x_j}} \tag{12}$$

The parameter matrix $W_r \in \mathbb{R}^{C\times(d_c+d_h))}$ and bias vector $b_r \in R_C$ are utilized in the final stage of sentiment analysis to classify the input data into different sentiment categories. Here, C represents the total number of sentiment categories. By using these tools, the proposed approach is capable of effectively classifying sentiment data into different categories with high accuracy.

3.2. **MAM Training.** The iterative update of parameter matrices and bias vectors in the proposed sentiment analysis approach is accomplished using backward propagation, a common technique in deep learning. In order to optimize the accuracy of the model, the cross entropy and regularization of all sentence classification results in the training set are utilized as the loss function. The relevant formulas are detailed in (13) and (14), providing a clear and concise approach to loss function calculation. Using this approach, the proposed model can efficiently and effectively update its parameters over time, resulting in more ac-curate sentiment classification. These findings have important implications for developing more effective sentiment analysis methods, and may be useful in various con-texts where sentiment information plays a critical role in decision-making processes.

$$J = -\sum_{i=1}^{C} g_i \log y_i + \lambda_r \left(\sum_{\phi\in\Phi} \phi^2\right) \tag{13}$$

$$\Phi = \Phi - \lambda_l \frac{\partial J(\Phi)}{\partial \Phi} \tag{14}$$

In this sentiment analysis approach, gi represents the true sentiment distribution in the reviews, while y_i corresponds to the model's prediction of the sentiment polarity for each review. The set of all parameters is denoted as Φ, while λ_r represents the parameter of the L2 regularization, and λ_l is the learning rate utilized during the parameter update process. By incorporating these elements into the model, the proposed approach can effectively and efficiently predict sentiment polarity in various contexts. Using L2 regularization and learning rate optimization ensures that the model remains stable over time, im-proving the accuracy and consistency of sentiment classification results. These findings may have important implications for developing more effective sentiment analysis methods, and may be helpful in various industries and applications where senti-ment information plays a critical role.

4. Experiment Results.

4.1. **Experiment Metric.** To assess the effectiveness of the proposed multilayered attention mechanism in sentiment analysis, this research undertook a meticulous sequence of tests encompassing diverse facets, including dataset selection, evaluation metrics, and hyperparameter adjustment. Performance evaluation of the model employed established metrics such as TP (true positive), TN (true negative), FP (false positive), and FN (false negative). The average accuracy was computed as the ratio of accurate predictions to the overall number of cases. Accuracy gauged the precision of positive predictions, while recall measured the proportion of accurately classified positive instances among all positive examples. Furthermore, the F1 score was employed as a comprehensive metric combining accuracy and coverage. These rigorous measurement standards were applied to appraise the model's performance. In this study, we aimed to assess the impact of a novel multilayer attention mechanism in a sentiment analysis model. To this end, a series of experiments were conducted, including selecting appropriate datasets, evaluation metrics, and hyper-parameter configurations. Specifically, the average accuracy and F1 score were computed

Table 1. Confusion Matrix.

		Predict	
		Positive	Negative
Actual	Positive	TP	FP
	Negative	FN	TN

using Equations (15) through (18) to evaluate the performance of the proposed model.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{15}$$

$$Precision = \frac{TP}{TP + FP} \tag{16}$$

$$Recall = \frac{TP}{TP + FN} \tag{17}$$

$$F1_{score} = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{18}$$

4.2. **Experiment Setup.** This investigation utilized the MOOC dataset sourced from a Chinese university to conduct an analysis of online course comments and gain insights into the emotional trends among students, with the objective of enhancing teaching quality and optimizing classroom efficiency. MOOC China University represents a prominent platform within China, offering a vast selection of over 10,000 online courses, including more than 1400 courses along with nationwide standing, in collaboration with 803 universities [42]. The dataset comprised a total of 11,293 online course reviews obtained from MOOCs in China, categorized into two distinct types of comments: positive and negative, encompassing 6164 positive comments and 5129 negative comments. To prepare the dataset for analysis, we performed several preprocessing steps to ensure the data quality and consistency. This included text normalization, removal of non-informative characters, and tokenization. We also applied natural language processing techniques such as stemming to reduce words to their root form and stop-word removal to eliminate common words that add noise to sentiment analysis. In this study, the dataset was divided into three subsets: a training set, a validation set, and a test set, allocated at ratios of 80%, 10%, and 10% respectively. Detailed distribution specifics of the dataset can be found in Figure 3 and Table 2. The main objective of this analysis was to assess the sentiment expressed in course comments and identify trends in sentiment patterns among senior students. The ultimate goal was to enhance the quality of teaching and promote student engagement.

Table 2. Dataset Detail.

Source	**Dataset**	**Positive**	**Negative**
MOOC China	Train	4609	4068
MOOC China	Validation	598	531
MOOC China	Test	600	630

In order to evaluate the sentiment analysis model applied to assessing online classroom feedback, a hierarchical attention mechanism is employed, and the evaluation is carried out using metrics such as average accuracy, precision, and recall. This evaluation aims to provide a comprehensive understanding of the strengths and weaknesses of the model. The F1 score is utilized to assess the performance of the MOOC model. Our experimental

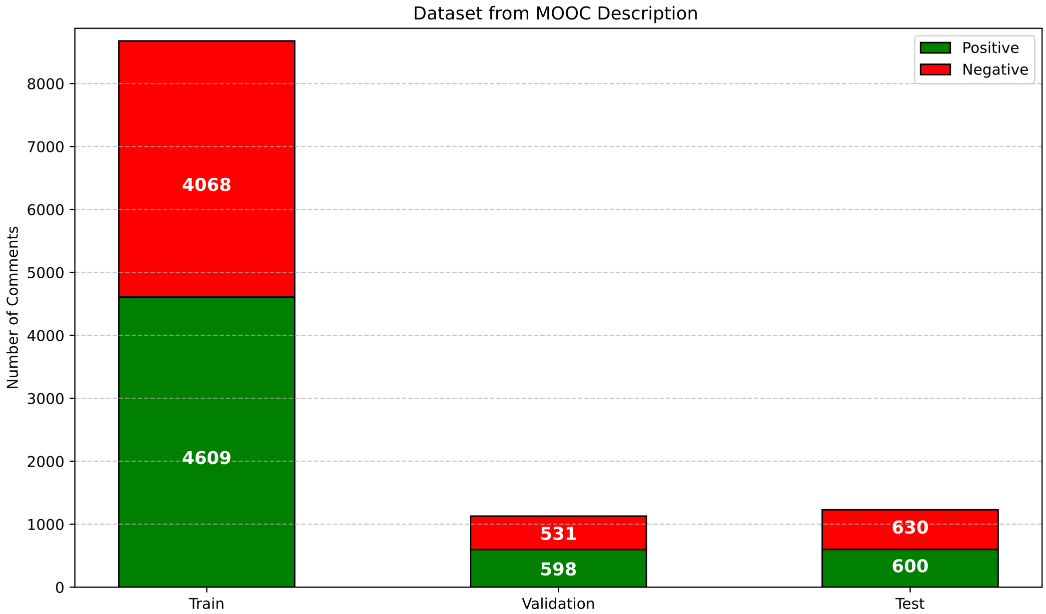

Figure 3. Dataset Description

configuration was carefully designed for clarity and reproducibility. We initialized the model using a single Chinese character, expanded through random vector generation to capture semantic depth. Key training parameters included the use of the Adam optimizer for its adaptive learning rate, a weight decay of 3e-5 for regularization, and a learning rate of 0.01, optimizing for both learning speed and stability. Our model's architecture featured 300 hidden layer units and CNN convolution windows of sizes 2, 3, and 4 to detect various semantic patterns. A dropout rate of 0.5 was set to prevent overfitting, promoting generalization. Inputs were tokenized text data, with the output as binary sentiment indicators.. The training parameters of the hierarchical attention mechanism are set as Table 3.

Table 3. Training parameters.

Parameter	**Value**
Optimizer	Adam
Weight decay	3e-5
Learning rate	0.01
Hidden Layer	300
CNN Convolution window	2, 3, 4
Dropout	0.5

4.3. **Experimental Results and Analysis.** The MAM model, as outlined in the proposal, underwent a training process spanning 30 epochs in order to attain a dependable level of accuracy. MAM also demonstrated appropriate loss levels throughout the training and validation stages of the study. The training and validation loss for the proposed model is depicted in Figure 4. Figure 4 exhibits a positive outcome in the model's learning curve, where the decrease in training and validation losses indicates effective learning. The plateauing of validation loss suggests a good juncture for early stopping, potentially optimizing the model's generalization abilities. This trend signifies that with some hyperparameter adjustments, the model is well-set to achieve an excellent balance in performance, illustrating the success of the chosen network architecture and training approach. The proposed MAM is utilized key evaluation measures, specifically Accuracy

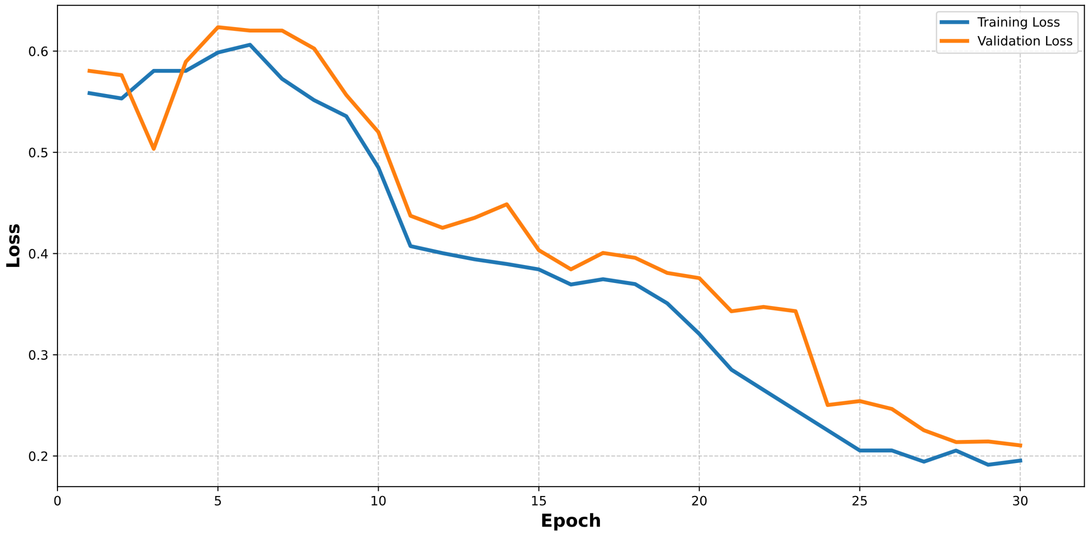

Figure 4. Training and validation loss of MAM

and Precision, to gauge its performance. During both the training and evaluation phases, the MAM model demonstrates exceptional effectiveness. In order to assess the efficacy of the MAM model, the author conducted a comparative analysis between the proposed model and four baseline models, with the aim of evaluating the respective outcomes. The proposed model exhibited a higher level of performance in comparison to other models in terms of accuracy . The accuracy of five methodologies is depicted in Figure 5. A metric

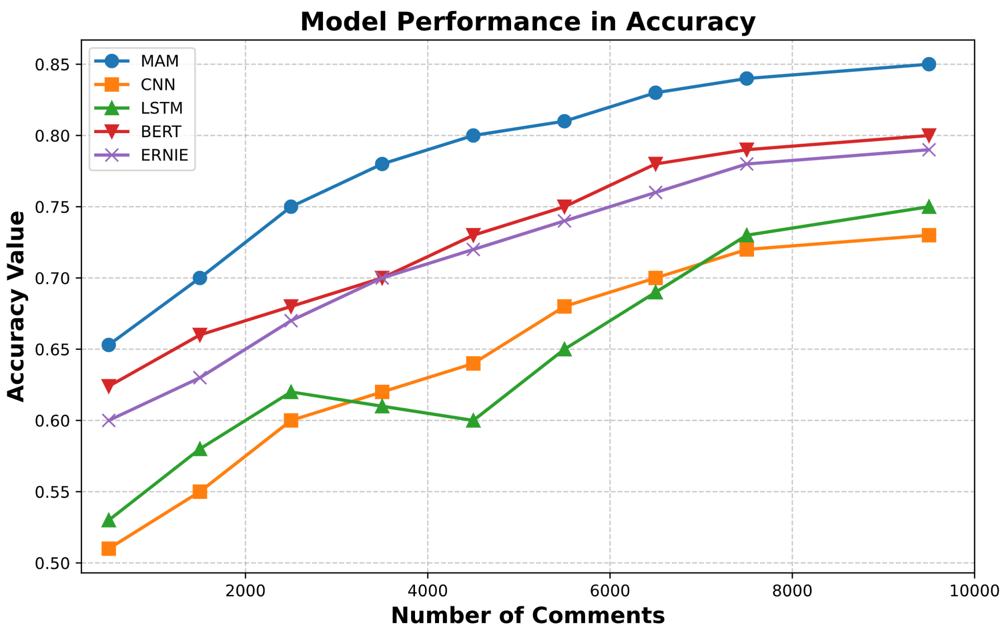

Figure 5. Accuracy of MAM and four baseline models

used to evaluate the efficacy of the learning model is precision, which is determined by dividing the number of TP by the total number of predictions made by the model. As precision calculating and analysis, the proposed model demonstrated superior performance compared to other models. Figure 6 illustrated the precision comparison among all the models analyzed. In addition to accuracy and precision, F1-score and recall are among

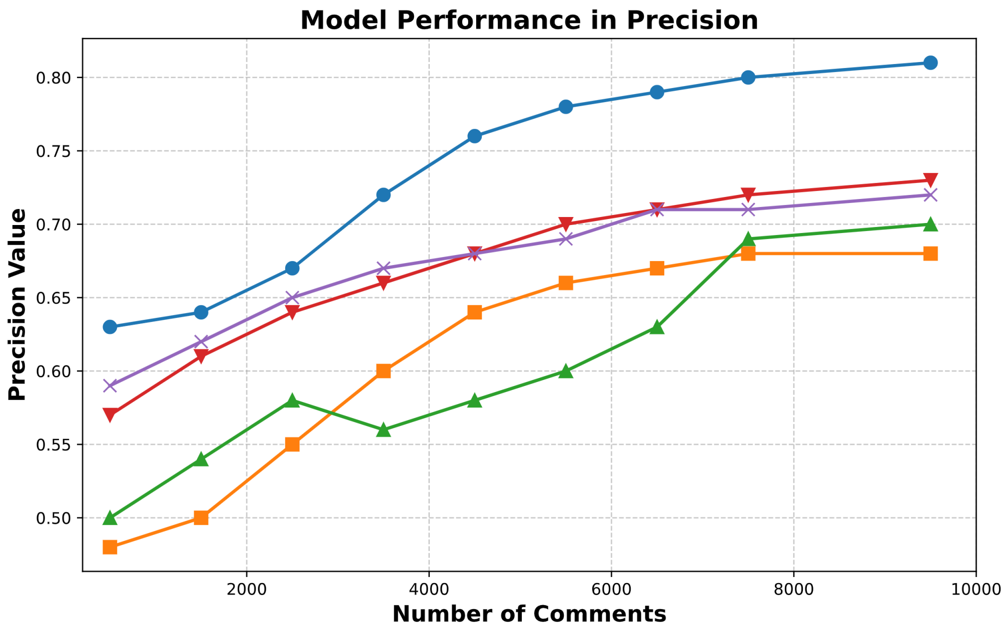

Figure 6. Precision of MAM and four baseline models

the evaluation metrics employed in the determination of the optimal model. Figure 7 and Table 3 illustrate the comparison of models across all evaluation metrics

Table 4. Comparative experimental results

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
MAM	0.85	0.81	0.92	0.88

The Table 4 outcomes suggest that the MAM approach exhibits superior performance in accuracy, precision, recall, and F1 score compared to the alternative models. MAM achieves an accuracy of 0.85, which is the highest among all the models. It also has the highest precision and recall scores, indicating that it is better at correctly iden-tifying positive and negative sentiment in the dataset. Additionally, the F1 score for MAM is 0.88, which is significantly higher than the other models. MAM is a model that amalgamates the advantageous features of CNN and LSTM models, enabling it to extract both local sentiment information and sequence information from text. In comparison to CNN, LSTM, BERT, and ERNIE, MAM exhibits superior performance. The MAM model incorporates gate and hierarchical attention mechanisms to augment its capacity to sift through local sentiment information by leveraging the overall text information acquired

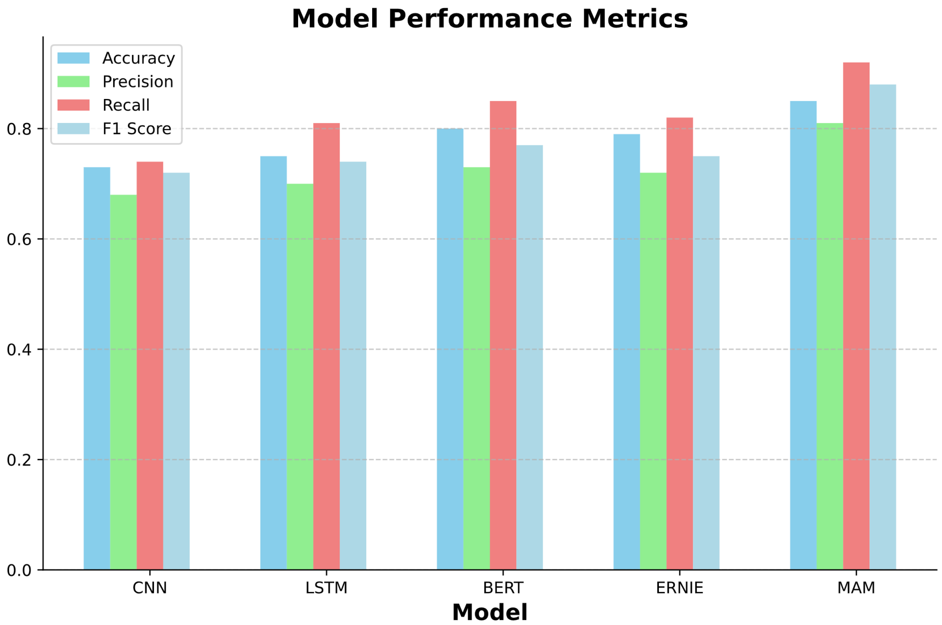

Figure 7. Performance of MAM and four baseline models

through LSTM. Additionally, it weighs the sentiment inclination of online classroom reviews. The aforementioned characteristics render MAM a robust contender for conducting sentiment analysis on virtual classroom evaluations.

4.4. **Comparing to The Other Related Works.** To conduct a comprehensive empirical assessment and analysis, the outcomes will be juxtaposed with contemporary research.Subsequently, an evaluation is conducted to assess the efficacy of five distinct recent models, namely our proposed MAM, SMA(Semantic Analysis model) [43], FeD (Feeling Distinguisher system) [44], and LSTM-Attention [40] , ASDA (Aspect-Based Sentiment Analysis) [45] (2022), TextCNN-CAT-FEW (Channel Attention TextCNN with Feature Word Extraction for Chinese Sentiment Analysis) [46] (2023). Table 5 depicts the comparision. The results suggest that the proposed MAM approach demonstrates superior

Table 5. Comparative experimental results with five recent models in literature.

Methods	Accuracy	Precision	Recall	F1
SMA [43]	-	0.76	0.94	0.84
FeD [44]	-	0.71	0.87	0.78
LSTM-Attention [40]	-	0.69	0.82	0.75
ASDA [45]	0.81	-	-	-
TextCNN-CAT-FEW [46]	0.83	0.81	0.80	0.80
MAM	0.85	0.81	0.92	0.88

performance compared to other recent models for sentiment analysis of online course reviews. MAM achieves the highest accuracy of 0.85, outperforming ASDA (0.81) and TextCNN-CAT-FEW (0.83). MAM also attains the highest precision, recall and F1 scores of 0.81, 0.92 and 0.88 respectively. Compared to SMA, FeD and LSTM-Attention, MAM shows significantly higher precision, indicating lower false positive rate in sentiment classification. While SMA and FeD exhibit high recall, their precision is considerably lower than

MAM, suggesting poorer ability to distinguish true positive sentiment. LSTM-Attention lags across all evaluation metrics. The robust results of MAM can be attributed to its technical novelty in integrating CNN and LSTM to extract multi-granularity features, enabling more accurate modeling of complex review texts. The gating mechanism selectively filters emotionally salient data, while hierarchical attention handles nuanced linguistic expressions.

The empirical comparisons validate MAM as a state-of-the-art advancement for sentiment analysis tailored to online course reviews. The results highlight its effectiveness in extracting localized and global sentiment features from texts and precisely classifying affective polarity. Further studies could assess MAM's applicability to other text categorization tasks and domains.

4.5. **Limitations and Future Work.** Despite the promising results demonstrated by the MAM, there are inherent limitations that need to be acknowledged and addressed in future research. One primary limitation lies in the model's dependency on the quality and diversity of the training dataset. The current model was trained and tested primarily on MOOC datasets, which may not fully represent the diverse linguistic expressions and sentiment constructs present in other online educational platforms or in different cultural contexts. This limitation could potentially affect the generalizability of the model to varied educational settings. Another challenge is the computational complexity associated with the hierarchical attention mechanism and LSTM components of the model, which might limit its scalability and real-time application. Optimizing the model for faster processing without compromising accuracy remains an area for future exploration.

Future work will focus on enhancing the model's robustness and applicability across different online educational platforms and languages. This includes extending the training datasets to encompass a wider range of linguistic styles and cultural contexts, thereby improving the model's adaptability and accuracy in diverse settings. Additionally, efforts will be made to streamline the model's architecture, potentially exploring more efficient algorithms or lighter neural network structures that maintain high performance while reducing computational load.

5. **Conclusion.** This research marks a significant advancement in sentiment analysis for online education through the development of a MAM. The MAM model stands out for its integration of CNN for local sentiment detection and LSTM for capturing deeper, hidden textual sentiments. A key innovation in MAM is its hierarchical attention mechanism, which effectively minimizes noise impact and emphasizes sentiments conveyed by key words, thereby enhancing judgment accuracy.Our experiments on MOOC datasets have demonstrated MAM's superior performance in sentiment analysis, highlighting its ability to extract nuanced sentiment information more accurately than traditional methods. Notably, MAM's gate mechanisms refine the sentiment extraction process, ensuring that only relevant affective information influences the model's output. Looking forward, we plan to enrich the word representation in MAM by incorporating additional data types such as speech, sentiment, and location contexts. This expansion aims to provide a more holistic view of sentiment in online course feedback. Additionally, we intend to collect and create more detailed attribute-level datasets for online courses. This effort will not only refine the accuracy of online course feedback analysis but also contribute to deeper engagement and interaction within online learning environments. These future developments promise to further elevate MAM's applicability and effectiveness in the evolving landscape of online education sentiment analysis.

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, Accessed: Apr. 19, 2023. [Online]. 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, pp. 289–298, 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, pp. 101–106, 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, p. 012002, Mar. 2021, doi: 10.1088/1742-6596/1860/1/012002.

[6] B. Su and J. Peng, "Sentiment Analysis of Comment Texts on Online Courses Based on Hierarchical Attention Mechanism," *Applied Sciences*, vol. 13, no. 7, Art. no. 7, Jan. 2023, doi: 10.3390/app13074204.

[7] B. Pang and L. Lee, "A sentimental education: sentiment analysis using subjectivity summarization based on minimum cuts," in *Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics*, in ACL '04. USA: Association for Computational Linguistics, July 2004, pp. 271-es. doi: 10.3115/1218955.1218990.

[8] J. Xiao, Z. Shao, S. Han, and Z. Li, "A Study on Sentiment Analysis of Online Course Review Information Based on XLNet-BiGRU," in *2022 12th International Conference on Information Technology in Medicine and Education (ITME)*, Oct. 2022, pp. 343–347. doi: 10.1109/ITME56794.2022.00081.

[9] 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, no. 2, Art. no. 2, Apr. 2023, doi: 10.24018/ejedu.2023.4.2.531.

[10] C. Alencia Hariyani, A. Nizar Hidayanto, N. Fitriah, Z. Abidin, and T. Wati, "Mining Student Feedback to Improve the Quality of Higher Education through Multi Label Classification, Sentiment Analysis, and Trend Topic," in *2019 4th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, Oct. 2019, pp. 359–364. doi: 10.1109/ICITISEE48480.2019.9003818.

[11] 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, Mar. 2022, doi: 10.1016/j.ijresmar.2021.10.011.

[12] F. Dalipi, K. Zdravkova, and F. Ahlgren, "Sentiment Analysis of Students' Feedback in MOOCs: A Systematic Literature Review," *Frontiers in Artificial Intelligence*, vol. 4, 2021, Accessed: Apr. 19, 2023. [Online]. Available: https://www.frontiersin.org/articles/10.3389/frai.2021.728708

[13] D. Bahdanau, K. H. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate: 3rd International Conference on Learning Representations, ICLR 2015," Jan. 2015.

[14] 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: Association for Computational Linguistics, Sep. 2015, pp. 1412–1421. doi: 10.18653/v1/D15-1166.

[15] Z. Yang, D. Yang, C. Dyer, X. He, A. Smola, and E. Hovy, "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: Association for Computational Linguistics, Jun. 2016, pp. 1480–1489. doi: 10.18653/v1/N16-1174.

[16] Y. Zhou, Q. Zhang, D. Wang, and X. Gu, "Text Sentiment Analysis Based on a New Hybrid Network Model," *Comput Intell Neurosci*, vol. 2022, p. 6774320, Dec. 2022, doi: 10.1155/2022/6774320.

[17] A. Chanaa and N. El Faddouli, "Sentiment Analysis on Massive Open Online Courses (MOOCs): Multi-Factor Analysis, and Machine Learning Approach," https://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/IJICTE.310004, Jan. 2017.

[18] S. Kumar, N. A. Jailani, A. R. Singh, and S. Panchal, "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.

[19] E. Cambria, A. Hussain, T. Durrani, and J. Zhang, "Towards a Chinese Common and Common Sense Knowledge Base for Sentiment Analysis," in *Advanced Research in Applied Artificial Intelligence*, H. Jiang, W. Ding, M. Ali, and X. Wu, Eds., in Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2012, pp. 437–446. doi: 10.1007/978-3-642-31087-4_46.

[20] K. Denecke and D. Reichenpfader, "Sentiment analysis of clinical narratives: A scoping review," *Journal of Biomedical Informatics*, vol. 140, p. 104336, Apr. 2023, doi: 10.1016/j.jbi.2023.104336.

[21] M. Z. Ansari, M. B. Aziz, M. O. Siddiqui, H. Mehra, and K. P. Singh, "Analysis of Political Sentiment Orientations on Twitter," *Procedia Computer Science*, vol. 167, pp. 1821–1828, Jan. 2020, doi: 10.1016/j.procs.2020.03.201.

[22] N. N. Reddy, E. Naresh, and V. Kumar B.P., "Predicting Stock Price Using Sentimental Analysis Through Twitter Data," in *2020 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT)*, Jul. 2020, pp. 1–5. doi: 10.1109/CONECCT50063.2020.9198494.

[23] A. Tsamados et al., "The ethics of algorithms: key problems and solutions," *AI & Soc*, vol. 37, no. 1, pp. 215–230, Mar. 2022, doi: 10.1007/s00146-021-01154-8.

[24] D. Ravi, C. Wong, F. Deligianni, M. Berthelot, J. A. Perez, B. Lo and G. Z. Yang , "Deep Learning for Health Informatics," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 1, pp. 4–21, Jan. 2017, doi: 10.1109/JBHI.2016.2636665.

[25] L. Yao, C. Mao, and Y. Luo, "Graph Convolutional Networks for Text Classification," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 01, Art. no. 01, Jul. 2019, doi: 10.1609/aaai.v33i01.33017370.

[26] R. Ragesh, S. Sellamanickam, A. Iyer, R. Bairi, and V. Lingam, "HeteGCN: Heterogeneous Graph Convolutional Networks for Text Classification," in *Proceedings of the 14th ACM International Conference on Web Search and Data Mining*, in WSDM '21. New York, NY, USA: Association for Computing Machinery, March 2021, pp. 860–868. doi: 10.1145/3437963.3441746.

[27] E. A. H. Khalil, Enas. M. F. El Houby, and H. K. Mohamed, "Deep Learning Approach in Sentiment Analysis: A Review," in *2020 15th International Conference on Computer Engineering and Systems (ICCES)*, Oct. 2020, pp. 1–10. doi: 10.1109/ICCES51560.2020.9334625.

[28] N. Yadati, M. Nimishakavi, P. Yadav, V. Nitin, A. Louis, and P. Talukdar, "HyperGCN: A New Method For Training Graph Convolutional Networks on Hypergraphs," in *Advances in Neural Information Processing Systems*, Curran Associates, Inc., 2019.

[29] X. Liu, X. You, X. Zhang, J. Wu, and P. Lv, "Tensor Graph Convolutional Networks for Text Classification," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 05, Art. no. 05, Apr. 2020, doi: 10.1609/aaai.v34i05.6359.

[30] 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, Art. no. 17, Jan. 2023, doi: 10.3390/math11173701.

[31] T.-Y. Wu, Q. Meng, Y.-C. Chen, S. Kumari, and C.-M. Chen, "Toward a Secure Smart-Home IoT Access Control Scheme Based on Home Registration Approach," *Mathematics*, vol. 11, no. 9, Art. no. 9, Jan. 2023, doi: 10.3390/math11092123.

[32] 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, no. 1, p. 36, Apr. 2023, doi: 10.1186/s13638-023-02245-4.

[33] C.-M. Chen, Z. Zhang, J. Wu, and K. Lakshmanna, "High Utility Periodic Frequent Pattern Mining in Multiple Sequences," *CMES*, vol. 137, no. 1, pp. 733–759, 2023, doi: 10.32604/cmes.2023.027463.

[34] C.-M. Chen, Y. Gong, and J. M.-T. Wu, "Impact of Technical Indicators and Leading Indicators on Stock Trends on the Internet of Things," *Wireless Communications and Mobile Computing*, vol. 2022, p. e9916310, May 2022, doi: 10.1155/2022/9916310.

[35] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "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: Human Language Technologies, Volume 1 (Long and Short Papers)*, Minneapolis, Minnesota: Association for Computational Linguistics, Jun. 2019, pp. 4171–4186. doi: 10.18653/v1/N19-1423.

[36] Z. Yang, Z. Dai, Y. Yang, J. Carbonell, R. R. Salakhutdinov, and Q. V. Le, "XLNet: Generalized Autoregressive Pretraining for Language Understanding," in *Advances in Neural Information Processing Systems*, Curran Associates, Inc., 2019.

[37] Y. Sun et al., "ERNIE 2.0: A Continual Pre-Training Framework for Language Understanding," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 05, Art. no. 05, Apr. 2020, doi: 10.1609/aaai.v34i05.6428.

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

[39] P. Liu, X. Qiu, and X. Huang, "Recurrent neural network for text classification with multi-task learning," in *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, in IJCAI'16. New York, New York, USA: AAAI Press, July 2016, pp. 2873–2879.

[40] Y. Wang, M. Huang, X. Zhu, and L. Zhao, "Attention-based LSTM for Aspect-level Sentiment Classification," in *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, Austin, Texas: Association for Computational Linguistics, Oct. 2016, pp. 606–615. doi: 10.18653/v1/D16-1058.

[41] J. Barnes, E. Velldal, and L. Øvrelid, "Improving sentiment analysis with multi-task learning of negation," *Natural Language Engineering*, vol. 27, no. 2, pp. 249–269, Mar. 2021, doi: 10.1017/S1351324920000510.

[42] L. Wang, G. Hu, and T. Zhou, "Semantic Analysis of Learners' Emotional Tendencies on Online MOOC Education," *Sustainability*, vol. 10, no. 6, Art. no. 6, Jun. 2018, doi: 10.3390/su10061921.

[43] 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, pp. 987–995, Jul. 2016, doi: 10.1109/JBHI.2016.2535721.

[44] 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: Association for Computational Linguistics, Jul. 2015, pp. 1556–1566. doi: 10.3115/v1/P15-1150.

[45] J. Melba Rosalind and S. Suguna, "Predicting Students' Satisfaction Towards Online Courses Using Aspect-Based Sentiment Analysis," in *Computer, Communication, and Signal Processing*, E. J. Neuhold, X. Fernando, J. Lu, S. Piramuthu, and A. Chandrabose, Eds., in IFIP Advances in Information and Communication Technology. Cham: Springer International Publishing, 2022, pp. 20–35. doi: 10.1007/978-3-031-11633-9_3.

[46] 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.