Centroid-Based Aspect Sentiment Analysis for MOOC Learner Reviews

Centroid-Based Aspect Sentiment Analysis for MOOC Learner Reviews

Naila Marir
Effat College of Engineering
Effat University, Jeddah, Saudi Arabia
nmarir@effat.edu.sa

Akila Sarirete
Effat College of Engineering
Effat University, Jeddah, Saudi Arabia
asarirete@effatuniversity.edu.sa

Tayeb Brahimi
Effat College of Engineering
Effat University, Jeddah, Saudi Arabia
tbrahimi@effatuniversity.edu.sa

2025 22nd International Learning and Technology Conference (L&T) | 979-8-3315-3221-5/25/$31.00 ©2025 IEEE | DOI: 10.1109/LT64002.2025.10940787

***Abstract*—Developing comprehensive analytics for Massive Open Online Courses (MOOCs) is essential for improving course design and enhancing learner engagement. In this work, we introduce MOOCSense, a multi-stage sentiment analysis module designed to analyze MOOC learner reviews and contribute to generating detailed MOOC analytics. In the first stage, we employ a mapping algorithm that extracts key MOOC-specific terms and central semantic phrases from the reviews. In the second stage, we propose a novel Centroid-Based Learning approach combined with the BERT (CLB) model to capture both implicit and explicit sentiment polarity in learner reviews, leveraging BERT's deep contextual understanding of natural language. By focusing on the central semantics of each review, our approach uncovers the emotional drivers behind learner engagement or dissatisfaction. This dual-stage module enables more accurate sentiment association with specific course aspects, enriching MOOC analytics with valuable insights. Experimental results demonstrate the effectiveness of our approach across various MOOC datasets, achieving an accuracy of 92%, making it a promising solution for generating in-depth learning analytics and supporting course improvement strategies.**

***Index Terms*—MOOC analytics, sentiment analysis, aspect-based sentiment analysis, BERT, centroid-based learning.**

I. Introduction

The digital revolution has profoundly transformed education, with online learning emerging as a cornerstone of modern educational practices [1]–[3]. In this landscape, the dynamic interaction between learners (humans) and digital platforms (machines) is crucial in creating responsive and personalized learning environments [4], [5]. Through AI and deep learning techniques, machines process and analyze learner feedback (sentiment), allowing systems to continuously adapt and tailor content to meet individual users' specific needs and preferences [5], [6]. This real-time adjustment enhances engagement by offering more personalized support to learners, thus improving their overall experience and fostering ongoing human-machine interaction.

The advent of MOOCs has marked a transformative era in education, offering scalable and accessible learning opportunities to a global audience [7], [8]. These platforms have democratized education, enabling learners worldwide to access diverse courses from renowned institutions. However, despite their vast potential, MOOCs face significant challenges, particularly in sustaining learner engagement and improving completion rates [9]. Addressing these issues requires innovative approaches to better understand learner behavior and enhance the educational experience.

One such promising approach is Sentiment Analysis (SA), a technique within Natural Language Processing (NLP) that systematically extracts and classifies sentiments, opinions, and emotions expressed in textual content [10], [11]. In the context of MOOCs, SA serves as a crucial tool for interpreting learner feedback, offering insights into engagement, satisfaction, and areas for improvement [12], [13]. The integration of Machine Learning (ML) and Deep Learning (DL) has further advanced the capabilities of SA [14], [15]. These methodologies allow for a deeper analysis of textual data, uncovering subtle sentiment patterns and capturing complex contextual nuances.

This study addresses the research question: *"How can DL-based sentiment analysis techniques be applied to interpret explicit and implicit sentiment patterns in MOOC learner feedback to enhance course sustainability and engagement?"* Using the Udemy dataset, we propose a centroid-based aspect sentiment analysis approach and evaluate several DL models for their effectiveness in capturing nuanced sentiment patterns. Our research contributes to the AI in education literature and offers practical insights for MOOC developers and educators to improve course design and learner experience.

The rest of the paper is organized as follows: Section 2 presents the State-of-the-Art, reviewing relevant literature and existing approaches in sentiment analysis for MOOCs. Section 3 details the Research Methodology. Section 4 provides the Results and Discussion, where we analyze the outcomes of the applied techniques and their implications for MOOC sustainability. Finally, Section 5 concludes the paper, summarizing key findings, acknowledging limitations, and suggesting future research directions.

II. Related Works

The integration of SA into MOOCs is crucial for enhancing educational platforms [12], [13]. However, as feedback data grows in complexity and volume, traditional ML techniques fall short, prompting a shift to DL methods. This section reviews existing DL models, focusing on their applications, limitations, and potential improvements.

An abundance of literature reviews attests the significance of sentiment analysis across different domains [16], [17]. These reviews typically categorize sentiment analysis techniques by methods such as ML or DL and the specific applications like social media or educational platforms. Traditional ML methods [18], [19], such as Support Vector Machines (SVM), Naive Bayes, and Artificial Neural Networks (ANN), have been extensively used to analyze sentiments in unstructured text, but they fall short when handling large-scale datasets like those generated by MOOCs [20], [21]. In contrast, DL models [15], [22], [23] such as Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNN), and Bidirectional Encoder Representations from Transformers (BERT) have demonstrated superior performance, particularly in detecting nuanced sentiment patterns in MOOCs. The literature confirms that DL techniques can automatically extract complex features from textual data, significantly improving sentiment classification accuracy and precision over traditional ML methods. Despite this advancement to the field of sentiment analysis, challenges remain in its application [10]. Talaat [24] proposed four deep four deep learning models by combining BERT with Bidirectional Long Short-Term Memory (BiLSTM) and Bidirectional Gated Recurrent Unit (BiGRU) algorithms. The research demonstrates that these hybrid models enhance accuracy in sentiment analysis tasks. However, these models rely heavily on extensive labeled data and excessive computation limiting their scalability to large-scale, unstructured datasets like MOOC reviews. Samy [12] introduced the BERT-LSTM-CNN (BLC), which integrates a pre-trained BERT base model for embedding, followed by a bi-directional LSTM and a CNN. The model effectively captures long dependencies and complex features, improving sentiment classification accuracy, however, it lacks a systematic mechanism for aspect-specific sentiment analysis. In another study, Sinha et al. [25] investigated the effectiveness of various deep learning techniques, including CNN, LSTM, LSTM-CNN, GRU, BERT, BERT-CNN, and BERT-LSTM. The study provides a comparative analysis highlighting the strengths of each model. However, its primary focus is on general sentiment classification rather than aspect-based sentiment analysis (ABSA). Additionally, the analysis does not explore the applicability of these models to domain-specific contexts, such as MOOCs, where feedback is often highly diverse and aspect-oriented.

Despite the success of DL in improving sentiment classification, more research is needed to assess how well these techniques generalize across different datasets and learning environments [21], [26]. The literature reveals critical gaps in the application of DL to sentiment analysis in MOOCs. One of the most notable is the limited research on aspect-based sentiment analysis (ABSA) within MOOCs, which could provide detailed insights into different aspects of the learning experience. Another gap is the insufficient focus on detecting implicit sentiment in MOOCs to capture nuanced learner feedback.

Addressing these challenges is essential for improving the personalization, learner engagement, and sustainability of MOOCs. This study builds upon these gaps to propose an innovative sentiment analysis framework tailored to the needs of MOOCs.

III. Research Methodology

In this section, we provide an overview of the foundational context, including the task formulation and fine-tuning process. Subsequently, we introduce our Aspect-Based Sentiment Analysis module, MOOCSense, which is systematically organized into four key phases: Input, Preprocessing, Sentiment Analysis, and Output. The core component, Sentiment Analysis, comprises two critical sub-stages—MOOC Aspect Extraction and MOOC Sentiment Detection. This modular structure of MOOCSense enables precise interpretation of learner feedback by identifying essential MOOC aspects and evaluating their associated sentiments, thereby offering actionable insights for enhancing MOOC design and delivery.

A. Background

1) Task Formulation: Our system aims to identify the sentiment (positive, negative, or neutral) expressed toward specific aspects of MOOC learner reviews, such as course content, instructor performance, platform usability, or learning experience, mentioned within each review. In this task, the input consists of a sentence $X = [x_1, x_2, \ldots, x_{n_X}]$, where each x_i represents a word, and n_X is the total number of words. Additionally, one or more aspect terms $a \in \{a_1, a_2, \ldots, a_{n_a}\}$, where the number of aspects in a sentence is denoted by n_a, are provided, with each aspect being a subsequence of the sentence. Our objective is to predict the sentiment polarity y for each aspect, where the possible sentiment labels $y \in Y$ are `POS` (positive), `NEG` (negative), or `NEU` (neutral).

2) Fine-Tuning BERT: BERT (Bidirectional Encoder Representations from Transformers) is a transformer-based language model known for its effectiveness across various natural language processing (NLP) tasks. For this study, we adopt the BERT-base model to address the Aspect-Based Sentiment Analysis (ABSA) task by framing it as a sentence classification problem. In the fine-tuning process, the input sentence $X = [x_1, x_2, \ldots, x_n]$ is passed through the BERT encoder, which transforms it into a sequence of hidden representations $H_X = [h_1, h_2, \ldots, h_n]$. A classification head is added on top of the BERT architecture to predict sentiment polarity for each identified aspect. Specifically, the embedding of the `[CLS]` token, which acts as a global representation of the input sequence, is fed into the classification layer. The model outputs a polarity prediction $Y = [y_1, y_2, \ldots, y_{n_a}]$, where each $y_i \in Y$ corresponds to a sentiment label (positive, negative, or neutral) associated with an aspect in the input. Fine-tuning BERT on ABSA enables the model to learn the task-specific details required to accurately classify the sentiment of multiple aspects within a given text.

B. Architecture Overview of the MOOCSense Module

The MOOCSense module is designed to provide comprehensive sentiment analysis of MOOC learner reviews through a structured four-phase process, as demonstrated in Figure 1: Input, Preprocessing, Sentiment Analysis, and Output. The sentiment analysis is further divided into two key sub-phases: MOOC aspect extraction and MOOC sentiment detection.

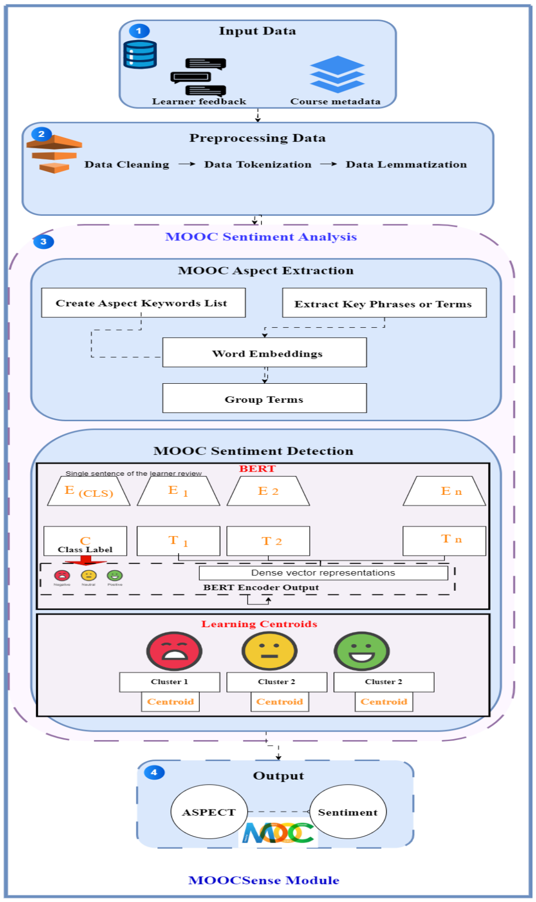

Fig. 1. Structural architecture of the proposed multi-layered system for aspect-based sentiment analysis.

1) Input Phase: The system first processes learner reviews, preparing them for further analysis. This step involves ingesting raw textual data from learner feedback.

2) Pre-processing Phase: In this phase, the input reviews undergo a series of transformations to standardize and normalize the data. Key steps include: Data Cleaning, Tokenization, Lemmatization and Data Annotation.

An example of the pre-processing step is as follows:

Original Learner Review:

"I really enjoyed this course! The content was informative and the instructor was engaging. However, I found some of the assignments to be too challenging."

- **Phase 1: Data cleaning** Output: "I really enjoyed this course The content was informative and the instructor was engaging However I found some of the assignments to be too challenging"
- **Phase 2: Data normalization** Output: "I really enjoyed this course the content was informative and the instructor was engaging however I found some of the assignments to be too challenging"
- **Phase 3: Data lemmatization** Output: "I really enjoy this course the content be informative and the instructor be engage however I find some of the assignment to be too challenging"
- **Phase 4: Data annotation** Output: [rating = 4, sentiment = "positive"]

3) Sentiment Analysis Phase: This phase is divided into two sub-phases:

MOOC Aspect Extraction: The system identifies key aspects of the MOOC, which are grouped into four broad categories:

- **Course:** Pertains to the course content and structure.
  Example: "The videos were easy to follow, and the syllabus was clear."
- **Instructor:** Refers to the instructor's knowledge and teaching abilities.
  Example: "The instructor explained complex concepts very clearly."
- **Assessment:** Focuses on quizzes, assignments, and other evaluations.
  Example: "The quizzes were well designed but a bit too hard."
- **Technology:** Relates to the platform's technical performance and user experience.
  Example: "The course videos kept buffering, which was frustrating."

The aspect extraction phase involves the following steps:

1) Creating aspect seed keywords for matching.
2) Preprocessing sentences to standardize them.
3) Extracting key phrases or terms associated with each aspect.
4) Utilizing word embeddings for semantic matching.
5) Grouping extracted terms under the most relevant aspect.

MOOC Sentiment Detection: Once aspects are extracted, the developed system applies a centroid-based learning method combined with BERT to determine the sentiment associated with each aspect. BERT embeddings help capture the contextual meaning of the review, while centroid-based learning groups the embeddings into clusters representing sentiment categories (positive, neutral, and negative).

4) Output Phase: The final phase provides a structured analysis, mapping each MOOC aspect to its respective sentiment (positive, neutral, or negative). This output offers valuable insights into learner feedback, facilitating improvements in MOOC design and delivery.

C. Centroid-Based Learning with BERT for Sentiment Detection

Centroid-Based Learning with BERT offers an effective approach to sentiment detection by leveraging contextual embeddings to identify representative centers for different sentiment classes. In this method, BERT converts learners' reviews into dense vector representations, capturing both semantic meaning and contextual nuances. These vectors are grouped into clusters based on their sentiment labels (e.g., positive, neutral, and negative), as illustrated in Algorithm 1.

Algorithm 1 Centroid-Based Learning with BERT for Sentiment Detection

Input: Learner reviews, Sentiment classes

Output: Sentiment for each aspect in the reviews

Input Preparation: Modify each review r to include aspect-specific masks:

$$r' = r + \text{[SSEP] sentiment of aspect: [mask]}$$

BERT Encoding: Pass the modified review r' through BERT to obtain embeddings E.

Centroid Learning: Initialize centroids $C_{\text{positive}}, C_{\text{neutral}}, C_{\text{negative}}$. **foreach** *masked embedding* E_{mask} **do**

  Assign E_{mask} to the nearest centroid C;

  Update centroids based on distances.

end

Mapping to Sentiment Labels: Assign labels based on nearest centroids:

$$\text{Label} = \begin{cases} \text{positive} & \text{if } E_{\text{mask}} \text{ is closest to } C_{\text{positive}} \\ \text{negative} & \text{if } E_{\text{mask}} \text{ is closest to } C_{\text{negative}} \\ \text{neutral} & \text{otherwise} \end{cases}$$

Training: Define loss functions:

$$\text{Total Loss} = \text{loss}_{\text{instance}} + \text{loss}_{\text{prototype}}$$

Prediction: For each review, retrieve embeddings and assign sentiment labels based on centroid proximity.

For each cluster, the centroid (average vector) is computed to represent the most central point for that sentiment class. During the classification phase, new reviews are encoded by BERT, and their vector representations are compared against the precomputed centroids. The system assigns the sentiment label of the closest centroid to the new review, enabling accurate classification even for nuanced or mixed sentiments. This centroid-based clustering method, enables scalable sentiment detection by reducing a large number of embeddings into a few representative centroids, significantly enhancing computational efficiency without compromising accuracy.

This design is particularly suited for aspect-based sentiment analysis, where quick yet reliable identification of sentiment is crucial to interpreting learner feedback and informing improvements in MOOC design and delivery.

IV. Results and Discussion

A. Dataset

The dataset used in this study originated from the Udemy MOOC platform and initially contained over **9 million comments** across **209,000 courses**. To ensure computational efficiency and retain high-quality data, we filtered this down to a representative subset of **104,712 entries**. This subset captures a diverse range of user experiences, with comments relevant to various aspects of MOOCs. The dataset was sourced from [27]. The dataset is characterized by several key attributes, which are detailed in Table I. These attributes are essential for the sentiment analysis process, as they provide both qualitative and quantitative data about the reviews.

The dataset was divided into **training**, **validation**, and **test** subsets to enable a structured model development pipeline.

TABLE I
Dataset Columns and Descriptions

Column	Description
id	Unique identifier for each review.
course_id	Identifier for each course, enabling course-based analysis.
date	Date the comment was posted, useful for temporal analysis.
display_name	An anonymized user identifier.
comment	The raw text of the review, serving as the primary data for sentiment analysis.
rate	A 1–5 numerical rating that reflects the user's satisfaction level and aids sentiment validation.

B. Experimental Setup

All experiments were conducted on a Google Colab environment with access to a high-performance GPU to accelerate model training and inference. For sentiment detection, the BERT base model was implemented using Hugging Face's Transformers library, consisting of 12 Transformer encoder layers with 12 attention heads per layer and a hidden size of 768, amounting to approximately 110 million parameters. Additionally, several other deep learning techniques, such as Long Short-Term Memory (LSTM) networks and Gated Recurrent Units (GRUs), were included for comparative analysis. These models were chosen for their effectiveness in capturing sequential patterns and managing long-term dependencies in text, making them suitable for identifying sentiment patterns in MOOC reviews.

C. Validation of the Proposed Approach

As detailed in Section III, our approach integrates aspect extraction and sentiment detection with Centroid-Based Learning using BERT to enhance MOOC analytics. The aspect extraction phase successfully identifies key MOOC aspects (e.g., Course, Instructor, Assessment, Technology), enabling a focused sentiment analysis. This is followed by Centroid-Based Learning with BERT in the sentiment detection phase, which clusters sentiment vectors to improve classification accuracy for positive, neutral, or negative sentiments. The aspect extraction performance, shown in Table II, underscores the method's precision in identifying relevant aspects within reviews.

Out of 104,712 reviews, 65,461 explicitly referenced these identified aspects (63%), while the remaining 37% involved other topics or implicit feedback not captured by predefined keywords. This diversity highlights that student feedback often contains implicit aspects, which were not fully captured by our approach. Future work will address this by refining aspect

TABLE II
PERFORMANCE OF ASPECT EXTRACTION AND SENTIMENT CLASSIFICATION ON MOOC REVIEWS

Aspect	Count	Positive	Neutral	Negative
Course	44133	38135	1445	4553
Instructor	10429	8465	485	1479
Assessment	7306	5000	603	1703
Technology	3593	2874	205	514
Total	**65461**	**54474**	**2738**	**8249**

extraction to include such implicit insights, providing a more comprehensive view of student perspectives.

The performance of the proposed sentiment detection approach is reported in Figure 2, where we compare the baseline BERT model with the Centroid-Based Learning (CBL) approach integrated with BERT.

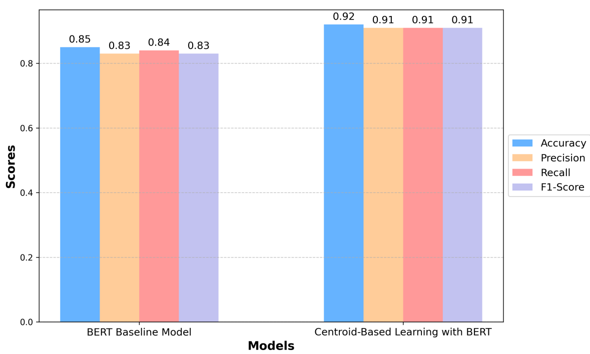

Fig. 2. Comparison of Sentiment Detection Models: BERT vs. BERT with Centroid Learning.

The comparison demonstrates a significant improvement in all evaluation metrics when using CBL with BERT. Specifically, the accuracy of the baseline BERT model was 85%, while the CBL-enhanced BERT model achieved an accuracy of 92%, indicating that CBL contributes to better generalization and more reliable sentiment detection. Similarly, precision increased from 83% with the baseline BERT model to 91% with CBL, reducing false positives and improving the accuracy of positive sentiment classification. Recall, which measures the model's ability to identify relevant sentiment instances, also saw an increase from 84% to 91%, suggesting that CBL helps capture more implicit sentiments in the data.

Moreover, the F1-score, which balances precision and recall, rose from 83% to 91%, indicating a more balanced and effective sentiment classification. These results show that the Centroid-Based Learning method enhances the ability of BERT to capture both explicit and implicit sentiment associations, leading to improved overall performance in sentiment detection.

D. Comparison with state-of-the-art methods

In this section, we compare the performance of our proposed approach against several state-of-the-art ML and DL models. The models compared include traditional ML algorithms such as Support Vector Machine (SVM), Neural Networks, K-Nearest Neighbors (KNN), Decision Trees, Random Forest, LightGBM, XGBoost, and AdaBoost, as well as deep learning models like Gated Recurrent Units (GRU) and Long Short-Term Memory (LSTM). Figure 3 summarizes the results of this comparison based on key evaluation metrics: accuracy, precision, recall, and F1-score.

As shown in Figure 3, our proposed approach, BERT with Centroid-Based Learning (CBL), outperforms all state-of-the-art methods across all metrics, achieving the highest scores in accuracy (92%), precision (91%), recall (91%), and F1-score (91%). Traditional models like SVM, K-Nearest Neighbors, and Decision Trees show lower performance, with accuracy ranging from 78% to 83%. Ensemble methods such as Random Forest and XGBoost perform better, with accuracies of 87% and 89%, respectively, but still fall short of our approach. Deep learning models like GRU and LSTM also demonstrate strong performance, with LSTM reaching 91% accuracy, but still slightly behind BERT with CBL. These results highlight that integrating Centroid-Based Learning with BERT enhances its ability to capture both explicit and implicit sentiment patterns, leading to superior performance in sentiment analysis tasks.

E. Case Study

The provided Table III showcases examples of use cases for MOOCSense. The system classifies the sentiment of reviews into three categories: positive, neutral, and negative and into two aspects: course or instructor. The examples demonstrate how the system can effectively classify the sentiment expressed in different review snippets.

TABLE III
MOOCSENSE USE CASE EXAMPLES

Comment	Aspects	Positive	Neutral	Negative
1. I just finished a MOOC on digital marketing and it was amazing!	Course	✓	-	-
2. The tutor's method of transmitting information varies from course to course.	Instructor	-	✓	-
3. This course is a waste of time. The content is poorly organized.	Course	-	-	✓

V. CONCLUSION

This research analyzed the application of deep learning-based sentiment analysis techniques to enhance the sustainability of MOOCs. The results show that deep learning models, particularly LSTM and GRU, outperform traditional machine learning models across all evaluation metrics. LSTM achieved the highest accuracy at 94%, while GRU followed closely with 92%, both demonstrating strong performance in detecting nuanced sentiment patterns in learner feedback. By leveraging sentiment data from learners, educational institutions can optimize course content and improve learner engagement, ultimately contributing to higher completion rates and more sustainable MOOC platforms.

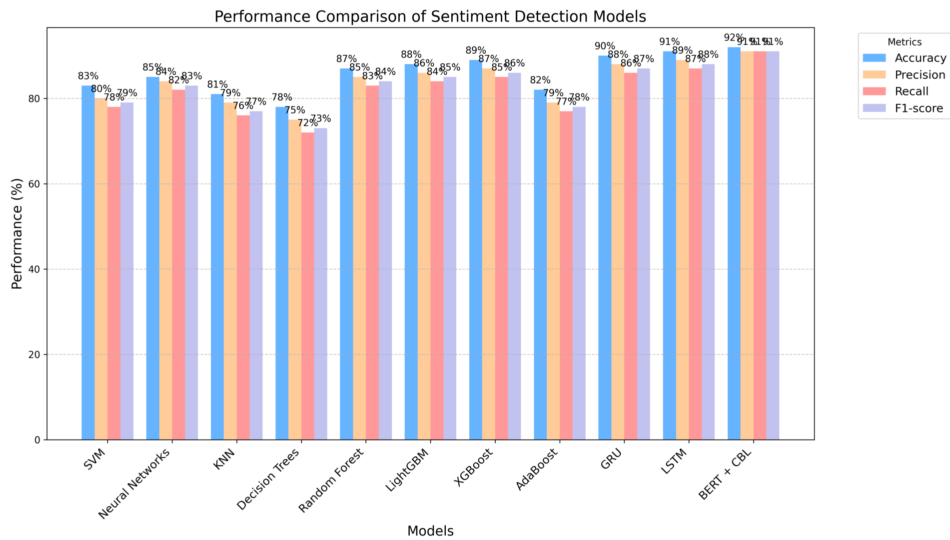

Fig. 3. Comparison with state-of-the-art methods.

Future work will focus on incorporating additional languages for sentiment analysis and exploring the integration of transfer learning techniques to enhance the model's adaptability across diverse educational contexts.

REFERENCES

[1] T. Brahimi and A. Sarirete. *Transforming Learning in STEAM: How AI Tools and Language Models Catalyze Educational Advancement*. Emerald Publishing, 2024.

[2] Nizar Bitar and Nitza Davidovich. Transforming pedagogy: The digital revolution in higher education. *Education Sciences*, 14:811, 2024.

[3] A. Sarirete, Z. Balfagih, and T. et al. Brahimi. Editorial - Artificial Intelligence: Towards Digital Transformation of Life, Work, and Education. *Procedia Computer Science*, 194:1–8, Dec 2021.

[4] Vicente Javier Clemente-Suárez. Navigating the new normal: Adapting online and distance learning in the post-pandemic era. *Education Sciences*, 14:19, 2024.

[5] Oyebola Olusola Ayeni, Nancy Mohd Al Hamad, and Onyebuchi Nneamaka et al. Chisom. Ai in education: A review of personalized learning and educational technology. *GSC Advanced Research and Reviews*, 18:261–271, 2024.

[6] Ivica Pesovski, Ricardo Santos, Roberto Henriques, and Vladimir Trajkovik. Generative ai for customizable learning experiences. *Sustainability*, 16:3034, 2024.

[7] T. Brahimi and A. Sarirete. Learning outside the classroom through MOOCs. *Computers in Human Behavior*, 51:604–609, 2015.

[8] Wiwit Ratnasari, Tzu Chou, and Chen Huang. From hype to reality: The changing landscape of mooc research. *Library Hi Tech*, 2024.

[9] Salvador Ceballos, Karla Maria Díaz López, and Paul Yitzen Ruelas Mexía. Conclusion and dropout of university students in mooc courses: analysis from the benefits perspective. *Práxis Educativa*, 19, 2024.

[10] Jamin Rahman Jim, Md Apon Riaz Talukder, and Malakar et al. Recent advancements and challenges of nlp-based sentiment analysis: A state-of-the-art review. *Natural Language Processing Journal*, page 100059, 2024.

[11] Robert Kasumba and Marion Neumman. Practical sentiment analysis for education: The power of student crowdsourcing. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 23110–23118, 2024.

[12] Adil Baqach and Amal Battou. A new sentiment analysis model to classify students' reviews on moocs. *Education and Information Technologies*, pages 1–28, 2024.

[13] Jack Fong. Sentiments and perspectives of academics about massive open online courses (moocs). 2024.

[14] Md Shofiqul Islam, Muhammad Nomani Kabir, and Ghani et al. Challenges and future in deep learning for sentiment analysis: a comprehensive review and a proposed novel hybrid approach. *Artificial Intelligence Review*, 57:62, 2024.

[15] Chinmayee Sahoo, Mayur Wankhade, and Singh et al. Sentiment analysis using deep learning techniques: a comprehensive review. *International Journal of Multimedia Information Retrieval*, 12:41, 2023.

[16] Mayur Wankhade, Chandra Sekhara, and Chaitanya A Kulkarni. A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, pages 5731–5780, 2022.

[17] Mika V. Mäntylä, Daniel Graziotin, and Miikka Kuutila. The Evolution of Sentiment Analysis - a Review of Research Topics, Venues, and Top Cited Papers. *Computer Science Review*, pages 16–32, 2018.

[18] Kian Long Tan, Chin Poo Lee, and Kian Ming Lim. A survey of sentiment analysis: Approaches, datasets, and future research. *Applied Sciences*, 13:4550, 2023.

[19] Arif Ullah, Sundas Naqeeb Khan, and Nazri Mohd Nawi. Review on sentiment analysis for text classification techniques from 2010 to 2021. *Multimedia Tools and Applications*, 82:8137–8193, 2023.

[20] Abdessamad Chanaa and Nour-eddine El Faddouli. Sentiment Analysis on Massive Open Online Courses (MOOCs). *International Journal of Information and Communication Technology Education (IJICTE)*, pages 1–22, 2022.

[21] Chenyu Qi. The challenges of online teacher education-sentiment analysis of course reviews. In *SHS Web of Conferences*, volume 190, page 01008. EDP Sciences, 2024.

[22] Neeraj Anand Sharma, ABM Shawkat Ali, and Muhammad Ashad Kabir. A review of sentiment analysis: tasks, applications, and deep learning techniques. *International Journal of Data Science and Analytics*, pages 1–38, 2024.

[23] Fatima Ahmed Al-Azazi and Mossa Ghurab. Ann-lstm: A deep learning model for early student performance prediction in mooc. *heliyon*, 9, 2023.

[24] Amira Samy Talaat. Sentiment analysis classification system using hybrid bert models. *Journal of Big Data*, 10:110, 2023.

[25] Swapnil Sinha, Abishek Jayan, and Rishabh Kumar. An analysis and comparison of deep-learning techniques and hybrid model for sentiment analysis for movie review. In *2022 3rd International Conference for Emerging Technology (INCET)*, pages 1–5. IEEE, 2022.

[26] Thanveer Shaik, Xiaohui Tao, Christopher Dann, Haoran Xie, Yan Li, and Linda Galligan. Sentiment analysis and opinion mining on educational data: A survey. *Natural Language Processing Journal*, 2:100003, 2023.

[27] M. G. Hossain. Udemy courses dataset. https://www.kaggle.com/datasets/hossaingh/udemy-courses, 2022. Accessed: 2024-10-26.