



ARTICLE

Enhancing Educational Materials: Integrating Emojis and AI Models into Learning Management Systems

Shaya A. Alshaya*

Computer Science Department, College of Science, Majmaah University, Al-Zulfi, 11932, Saudi Arabia

*Corresponding Author: Shaya A. Alshaya. Email: shaya@mu.edu.sa

Received: 17 December 2024; Accepted: 03 March 2025; Published: 16 April 2025

ABSTRACT: The integration of visual elements, such as emojis, into educational content represents a promising approach to enhancing student engagement and comprehension. However, existing efforts in emoji integration often lack systematic frameworks capable of addressing the contextual and pedagogical nuances required for effective implementation. This paper introduces a novel framework that combines Data-Driven Error-Correcting Output Codes (DECOC), Long Short-Term Memory (LSTM) networks, and Multi-Layer Deep Neural Networks (ML-DNN) to identify optimal emoji placements within computer science course materials. The originality of the proposed system lies in its ability to leverage sentiment analysis techniques and contextual embeddings to align emoji recommendations with both the emotional tone and learning objectives of course content. A meticulously annotated dataset, comprising diverse topics in computer science, was developed to train and validate the model, ensuring its applicability across a wide range of educational contexts. Comprehensive validation demonstrated the system's superior performance, achieving an accuracy of 92.4%, precision of 90.7%, recall of 89.3%, and an F1-score of 90.0%. Comparative analysis with baseline models and related works confirms the model's ability to outperform existing approaches in balancing accuracy, relevance, and contextual appropriateness. Beyond its technical advancements, this framework offers practical benefits for educators by providing an Artificial Intelligence-assisted (AI-assisted) tool that facilitates personalized content adaptation based on student sentiment and engagement patterns. By automating the identification of appropriate emoji placements, teachers can enhance digital course materials with minimal effort, improving the clarity of complex concepts and fostering an emotionally supportive learning environment. This paper contributes to the emerging field of AI-enhanced education by addressing critical gaps in personalized content delivery and pedagogical support. Its findings highlight the transformative potential of integrating AI-driven emoji placement systems into educational materials, offering an innovative tool for fostering student engagement and enhancing learning outcomes. The proposed framework establishes a foundation for future advancements in the visual augmentation of educational resources, emphasizing scalability and adaptability for broader applications in e-learning.

KEYWORDS: Emoji Integration; artificial intelligence in education; learning management systems; educational materials enhancement; student engagement

1 Introduction

The rapid evolution of digital technologies has dramatically reshaped education, leading to a growing reliance on e-learning systems within educational institutions and corporate training programs. The demand for these systems continues to surge, driven by the need for flexible, scalable, and accessible solutions that cater to a diverse range of learners. E-learning systems enable both students and professionals to engage with content at their own pace, thereby enhancing knowledge retention and adapting to individual learning



preferences. This transformation is not only present in academic settings but also in professional environments, where training programs are increasingly delivered online to meet the evolving skill requirements of modern workplaces. Learning Management Systems (LMS) have thus become vital in supporting digital-first approaches to education, offering insights into learner progress and facilitating personalized learning pathways [1,2].

To enhance learning performance, several methods have been incorporated into LMS platforms, each tailored to optimize engagement, comprehension, and retention. Among these techniques are gamification, which uses points, badges, and other game-like elements to motivate learners, and adaptive learning, which employs artificial intelligence to deliver content suited to individual performance [3]. Blended learning, combining online and face-to-face instruction, supports hybrid models that reinforce learning through varied modes. Microlearning, which presents information in short, focused segments, is particularly useful for skill-based or technical education, while social learning within LMS platforms encourages peer interaction, fostering collaboration and shared learning experiences. Additionally, analytics-driven insights enable data-informed adaptations to course materials, and the integration of interactive elements, such as quizzes and drag-and-drop tasks, facilitates hands-on engagement [4]. Recently, some systems have started employing sentiment analysis tools to monitor learner emotions and adapt content based on engagement and motivation [5].

While many of these techniques focus on cognitive engagement, the role of visual and emotional cues, specifically emojis, in digital learning has attracted growing attention. Previous studies have shown that emojis in digital communication can enhance emotional expression, clarify intent, and make text-based interactions more engaging. Within e-learning, research has explored the role of emojis in various contexts, such as feedback, peer communication, and course materials [6]. Studies suggest that emojis in feedback can enhance clarity and motivation, creating a supportive learning atmosphere. Additionally, the use of emojis in course materials has been shown to reduce cognitive load by presenting information in visually appealing ways, helping learners grasp complex concepts with relative ease [7]. In discussion forums, emojis have been found to promote interaction by allowing learners to express emotions, thereby fostering a sense of community within digital classrooms [8]. However, in the context of computer science education, students often face challenges in understanding abstract and complex concepts, such as algorithms, data structures, and programming logic. Traditional text-based explanations can sometimes be insufficient in conveying intricate problem-solving strategies or debugging processes effectively. The integration of emojis into computer science learning materials presents an opportunity to address these challenges by visually reinforcing key concepts, signaling important patterns, and providing intuitive cues to aid comprehension.

This paper addresses these challenges by proposing a novel framework that leverages advanced AI techniques to optimize emoji placement within course materials. The hybrid architecture combines DECOC, LSTM networks, and ML-DNN to analyze the emotional tone, context, and complexity of educational content, enabling precise and effective emoji integration. Unlike previous models that rely on heuristic-based approaches, the proposed framework dynamically adapts to content nuances, ensuring that emojis enhance rather than distract from learning objectives.

The originality of this research lies in the development of the EduEmoji-CS dataset, an extensive resource comprising annotated computer science materials, and the application of advanced AI methodologies to systematically address emoji placement. This dataset, paired with the hybrid model, enables the exploration of complex relationships between content, emotion, and pedagogy. The contributions of this paper are threefold: first, it introduces a scalable AI-driven solution to enhance learning materials through emoji integration; second, it provides empirical evidence of the model's efficacy through rigorous validation; and third, it establishes a foundation for future advancements in personalized learning environments. This

research not only bridges significant gaps in educational content enhancement but also offers practical applications for improving engagement and comprehension in e-learning systems globally.

This paper aims to investigate the impact of emojis on learner understanding and engagement within e-learning environments. Specifically, the research seeks to explore the integration of emojis across different sections of e-learning courses, facilitated by an AI-driven model that selects appropriate emojis based on course content and learner needs. By employing artificial intelligence to dynamically determine emoji placement, this approach aims to ensure that each emoji is contextually relevant, enhancing comprehension and supporting a high level of understanding. The purpose is to provide empirical insights into the effectiveness of emojis as a tool to enhance digital learning, offering evidence of their potential to improve the quality of educational experiences in online settings. Through this study, we aim to contribute to the broader understanding of how emojis can support cognitive and emotional engagement in e-learning, ultimately driving more impactful learning outcomes.

The paper is organized as follows: [Section 2](#) reviews related work on emoji integration in educational content and the application of AI models for contextual and sentiment analysis. [Section 3](#) introduces the proposed framework, detailing the hybrid DECOC-LSTM-ML-DNN architecture and its alignment with pedagogical objectives. [Section 4](#) presents the survey design for evaluating emoji integration in educational materials, including methodology, participant demographics, and analysis approaches. [Section 5](#) discusses the implementation process, including dataset development, preprocessing, and model training, followed by results and discussion. Finally, [Section 6](#) concludes the study and provides insights into future research directions to enhance educational content through AI-driven solutions.

2 Related Works

The integration of emojis into e-learning platforms has been the subject of various studies, each exploring their potential to enhance communication, engagement, and learning outcomes. This section reviews notable research in this domain, highlighting their methodologies, findings, and implications.

Sia et al. [9] investigated the role of emojis in text-based electronic-mediated communication (TEMC) within higher education. Their study employed the theories of source credibility and stimulus-organism-response to develop a model assessing how emoji usage influences perceptions of instructor credibility, learning motivation, and performance. Data collected from 348 students indicated that emojis positively affect perceptions of instructor expertise, trustworthiness, and likability, which in turn enhance student motivation and performance.

Blobstein et al. [10] analyzed the use of instructor-selected emojis during online discussions in e-learning settings. Their study found that emojis facilitated emotional expression and increased student engagement, contributing to a more interactive and supportive learning environment. The authors suggest that strategically incorporating emojis can bridge the emotional gap often present in digital communication, thereby enhancing the overall learning experience.

Mokhamed et al. [11] conducted a comparative analysis of machine learning and deep learning algorithms for emoji prediction in Arabic text. While not directly focused on e-learning, their findings offer valuable insights into how predictive models can enhance communication within educational platforms. The study demonstrated that advanced algorithms could accurately predict appropriate emojis based on textual context, suggesting potential applications in developing intelligent e-learning systems that adapt to learners' emotional cues.

Zarkadoulas et al. [12] conducted an empirical study on the use of emojis to express emotions in digital mathematics classes. Their research involved a questionnaire survey among 100 students, who selected

emojis reflecting their feelings when faced with mathematical problems. The results indicated that emojis effectively conveyed students' emotional states, providing instructors with valuable insights to tailor their teaching strategies accordingly. This study highlights the utility of emojis as tools for emotional expression and feedback in e-learning environments.

Bai et al. [13] conducted a systematic review examining the development, usage, function, and application of emojis in computer-mediated communication, including educational contexts. Their review highlighted the potential of emojis to enhance emotional expression and engagement in e-learning environments. The authors emphasized the need for further research to explore the specific applications and impacts of emojis within educational settings, particularly concerning learning outcomes and student engagement.

Table 1 summarizes the shortcomings of the above related works on emoji integration in education.

Table 1: Limitations in existing literature on emoji integration in education

Study	Focus area	Identified shortcomings	Implications
Sia et al. (2024) [9]	Emoji use in text-based communication for higher education	Limited focus on contextual relevance and pedagogical alignment	Existing approaches do not ensure that emojis align with learning objectives, reducing their effectiveness in enhancing comprehension
Blobstein et al. (2024) [10]	Instructor-selected emojis in online discussions	Lack of automation in emoji placement	Requires manual selection, making it time-consuming and inconsistent across different learning materials
Mokhamed et al. (2024) [11]	Machine learning models for emoji prediction in Arabic text	Absence of domain-specific applications for educational contexts	Existing models are trained on general text data, making them unsuitable for structured learning content like computer science courses
Zarkadoulas et al. (2024) [12]	Emotional expression using emojis in mathematics education	Limited to self-reported emotional feedback rather than pedagogical impact	Lacks empirical validation of emoji effectiveness in improving learning outcomes beyond emotional engagement
Bai et al. (2019) [13]	Systematic review of emoji applications in communication and education	No standardized framework for integrating emojis into course materials	The absence of structured methodologies results in inconsistent and arbitrary emoji use across educational platforms
Eisner et al. (2016) [14]	Emoji representations in sentiment analysis	Does not consider learning-specific contexts in emoji use	Most studies focus on sentiment and opinion mining rather than instructional applications
Shiri et al. (2023) [15]	Comparative analysis of deep learning models for Natural language processing (NLP)	Lack of multi-layered AI frameworks for contextualized emoji recommendations	Existing models fail to incorporate pedagogical insights for precise emoji placement in educational settings

[Table 2](#) summarizes the key aspects of the reviewed studies, highlighting their methodologies, findings, and limitations.

Table 2: Comparative analysis of emoji integration studies

Demographic category	Subcategory	Satisfaction with emoji-enhanced materials	Increased engagement	Improved understanding	Appropriate emoji placement	Willingness to use in future
Gender	Male	80%	75%	70%	85%	83%
	Female	90%	85%	80%	95%	95%
Academic level	First year	88%	84%	76%	92%	90%
	Second year	85%	80%	75%	90%	88%
	Third year	82%	78%	72%	88%	85%
	Fourth year	80%	75%	70%	85%	83%
Age group	18–20 years	90%	85%	80%	95%	95%
	21–23 years	85%	80%	75%	90%	88%
	24–26 years	80%	75%	70%	85%	83%
Ethnicity	Saudi	85%	80%	75%	90%	88%
	Non-Saudi	83%	78%	73%	88%	85%
	Arab					
	Other	80%	75%	70%	85%	83%

In light of the reviewed studies, it is imperative to propose an innovative approach to enhance learning efficiency and effectiveness. Building upon the insights from previous studies, we propose an AI-driven model utilizing DECOC-LSTM-ML-DNN architectures to identify optimal emoji placement within computer science courses. This model aims to enhance learner engagement and comprehension by strategically integrating emojis that align with course content and learner needs. The effectiveness of this approach will be evaluated through surveys assessing the impact of emojis on learning outcomes, providing empirical evidence to support the integration of emojis in e-learning environments.

3 Proposed AI-Driven Model for Emoji Integration in Learning Content

To enhance learner engagement and comprehension in computer science courses through the strategic integration of emojis, we propose a multi-phase approach leveraging advanced AI techniques as shown in [Fig. 1](#).

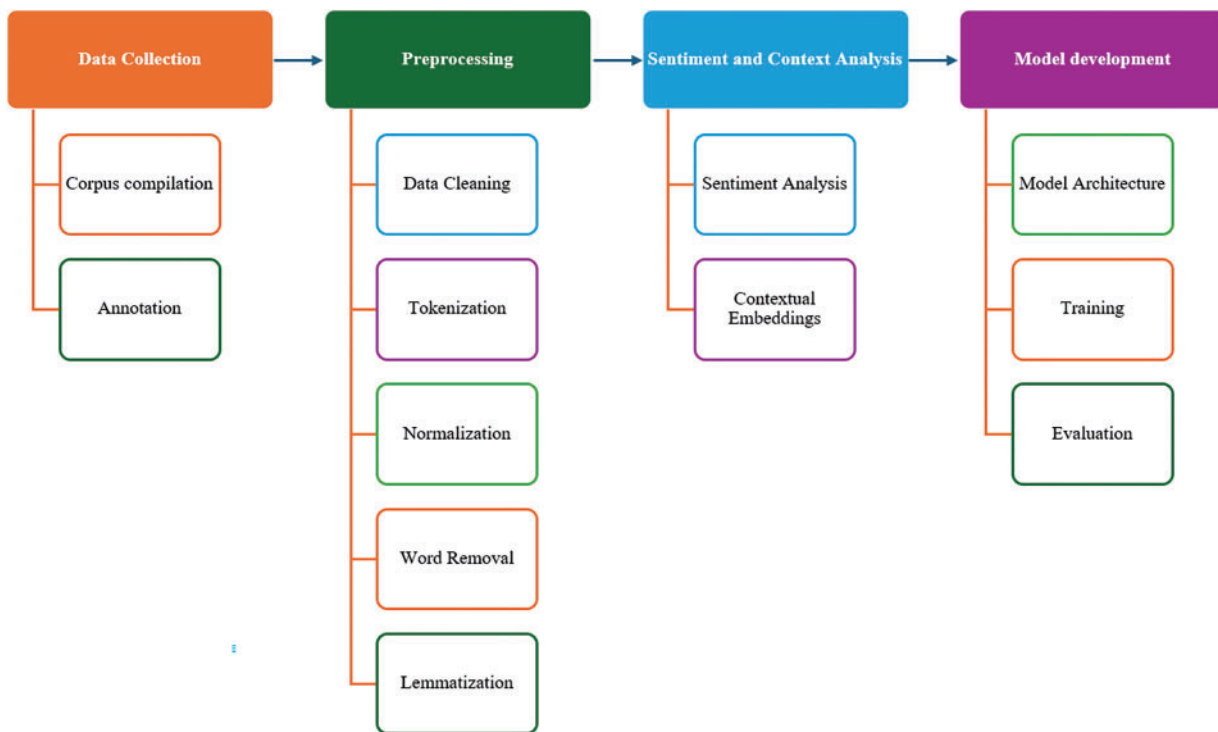


Figure 1: Overview of the proposed emoji integration workflow

3.1 Data Collection

This phase aims to enhance comprehension and engagement in computer science education through the strategic integration of contextually appropriate emojis into course materials. The methodology involves compiling a comprehensive dataset that includes lecture notes, presentation slides, assignments, and discussion forum entries, thereby providing a holistic view of the educational content and interactions within the course. Subject matter experts will annotate this dataset to identify sections where the integration of emojis could potentially enhance comprehension or increase student engagement. By analyzing students' feedback derived from these materials, the study seeks to pinpoint specific content segments that present challenges to learners. Incorporating emojis into these sections is anticipated to facilitate better understanding and retention of complex concepts, leveraging the communicative power of emojis to create a more engaging and supportive learning environment.

The developed annotated dataset untitled “EduEmoji-CS” is structured as follows:

- Document ID: A unique identifier assigned to each document to facilitate tracking and referencing.
- Document Type: Categorization of the material, such as lecture notes, presentation slide, assignment, or survey entry.
- Course Information: This encompasses the official course title, the name(s) of the instructor(s), the educational institution offering the course, and the academic term during which the course was conducted.
- Content: The full textual content of the document, along with references to associated media like images, videos, or audio files.
- Annotations: Each annotated section within the document has a unique Section ID. The specific portion of text being annotated (Text Segment) is identified, along with the nature of the annotation (e.g., concept explanation, example, question, or student feedback). An assessment of the complexity of the

text segment, categorized as low, medium, or high (Difficulty Level), is provided. Suggested emojis to enhance comprehension or engagement for the annotated segment is included, accompanied by a brief explanation justifying their choice and detailing how they are expected to aid understanding or engagement.

- **Metadata:** Information about the individual(s) who created the document (Author), the original creation date (Date Created), the date when the annotation was performed (Date Annotated), and the individual(s) responsible for the annotation (Annotator) is recorded.
- **Contextual Information:** This includes the background knowledge required to understand the content (Prerequisite Knowledge), the educational goals associated with the content (Learning Objectives), and citations or links to supplementary materials (References).

Table 3 presents an example of entry from the EduEmoji-CS Dataset following the description above.

Table 3: Annotation example for the introduction to data structures course

Field	Annotation details
Document ID	CS101_Lecture_05
Document Type	Lecture Notes
Course Information	Course Title: Introduction to Data Structures Instructor: Dr. XXXXX Institution: Majmaah University Semester: Spring 2024
Content	“In a binary search tree (BST), each node has at most two children, where the left child contains values smaller than the parent node, and the right child contains values greater than the parent node. This structure allows for efficient searching, insertion, and deletion operations, typically in $O(\log n)$ time complexity. However, if the tree becomes unbalanced, search performance may degrade to $O(n)$. To maintain balance, self-balancing BSTs like Adelson-Velsky and Landis (AVL) and Red-Black Trees are used.”
Annotations	
Section ID	CS101_Lec05_Slide03
Text Segment	“However, if the tree becomes unbalanced, search performance may degrade to $O(n)$.”
Annotation Type	Concept Explanation
Difficulty Level	Medium
Suggested Emoji	⚠️ (Warning Sign)
Justification	The warning sign emoji emphasizes the potential performance issue of unbalanced trees, alerting students to an important limitation of BSTs.
Metadata	
Author	Dr. XXXXX
Date Created	05 March 2024
Date Annotated	10 March 2024
Annotator	AI Model (EduEmoji-CS)
Contextual Information	

(Continued)

Table 3 (continued)

Field	Annotation details
Prerequisite Knowledge	Basic understanding of binary trees and recursion
Learning Objectives	- Understand the structure and operations of a binary search tree (BST). - Recognize the performance implications of unbalanced trees. - Learn about self-balancing tree techniques.
References	“Data Structures & Algorithms in Python,” 2nd Edition, Chapter 6

This proposed structured dataset ensures comprehensive documentation and facilitates systematic analysis of the dataset. Such organization supports the development of models aimed at enhancing educational materials through the strategic integration of emojis, thereby improving learner engagement and comprehension. In this paper, we are focusing in six commonly used emojis which is described in the [Table 4](#).

Table 4: Used emojis in the EduEmoji-CS dataset

Emoji	Name	Typical meaning	Sample
😊	Smiling Face with Smiling Eyes	Indicates a positive outcome or successful completion.	“Congratulations on successfully implementing the sorting algorithm! 😊”
😭	Crying Face	Highlights common pitfalls or errors to avoid.	“Be cautious with off-by-one errors in loops; they can lead to unexpected results. 😭”
😬	Pouting Face	Emphasizes critical warnings or prohibited actions.	“Plagiarism is a serious violation of academic integrity and will not be tolerated. 😬”
😎	Smiling Face with Sunglasses	Denotes advanced topics or best practices.	“Utilizing design patterns can make your code more efficient and maintainable. 😎”
🤔	Thinking Face	Introduces thought-provoking questions or prompts critical thinking.	“Consider: How does the choice of data structure impact algorithm efficiency? 🤔”
😴	Sleeping Face	Indicates less critical information or sections that can be reviewed later.	“The detailed history of programming languages is available in the appendix. 😴”

3.2 Preprocessing

Preprocessing is a critical phase in preparing annotated text data for model training, ensuring that the data is clean, consistent, and structured appropriately for analysis. This process involves several key tasks. First, data cleaning focuses on removing irrelevant information that could introduce noise into the dataset.

For instance, extraneous symbols, HyperText Markup Language (HTML) tags, or unrelated metadata present in lecture notes or discussion forum entries are eliminated. Next, tokenization involves breaking down the cleaned text into individual units, such as words or phrases, known as tokens. Tokenization facilitates the analysis of text at a granular level and ensuring consistency in formatting and structure is vital for accurate analysis. Normalization processes include converting all text to lowercase to maintain uniformity and handling contractions or abbreviations. This transformation ensures that different forms of a word or phrase are treated equivalently during analysis. Additionally, special characters and punctuation that do not contribute to meaning are removed to prevent them from affecting text embeddings and model training. Stopword removal ensures that only semantically relevant terms are retained, improving computational efficiency and reducing noise in text representation. Named entity recognition (NER) techniques can also be applied to preserve domain-specific terms, preventing the loss of critical information in technical course materials. Then, commonly used words that do not contribute significant meaning to the text, such as “is,” “and,” or “the,” are removed to focus on the more informative parts of the text in the word removal process. This step helps in reducing the dimensionality of the data and highlights the essential terms. Finally, lemmatization reduces words to their base or root form, known as a lemma, which helps in treating different forms of a word as a single item. By lemmatizing, the analysis can focus on the core meaning of words without being affected by their various inflections.

Implementing these preprocessing steps is essential to facilitate accurate and efficient analysis in the subsequent phases of the research. They ensure that the text data is in a standardized format, free from noise, and ready for effective model training.

3.3 Sentiment and Context Analysis

In the sentiment analysis phase, Natural Language Processing (NLP) techniques are employed to assess the emotional tone of educational content, identifying areas where emojis can effectively convey or reinforce sentiment [16]. This process involves the use of advanced models, such as Bidirectional Encoder Representations from Transformers (BERT) [17], to generate contextual embeddings that capture the nuanced meanings of text segments, thereby informing the strategic selection of appropriate emojis [18].

These techniques, also known as opinion mining, is a subfield of NLP that focuses on determining the sentiment expressed in a piece of text, categorizing it as positive, negative, or neutral. Traditional methods often rely on machine learning algorithms trained on labeled datasets to identify sentiment based on word usage patterns [19]. However, these approaches may struggle with context-dependent meanings and the subtleties of human language.

To address these challenges, this attempt utilizes BERT, a transformer-based model that generates contextualized word embeddings. Unlike static embeddings, BERT captures the meaning of words in relation to their surrounding context, allowing for a more accurate representation of language nuances. BERT’s bidirectional training enables it to consider the entire sentence, effectively disambiguating such words. By applying BERT to the educational corpus, the model generates embeddings that reflect the contextual meaning of text segments. These embeddings are then analyzed to detect underlying sentiments, such as frustration, confusion, or enthusiasm, which may not be explicitly stated.

The insights gained from sentiment analysis guide the selection of emojis that align with the identified emotional tones. In confusion statement, for example, an emoji representing confusion (e.g., 🤔) could be added alongside the text to acknowledge the student’s struggle and provide a visual cue that the material may be challenging. This empathetic approach can enhance student engagement and foster a supportive learning environment.

Therefore, integrating sentiment analysis with contextual embeddings allows for a nuanced understanding of educational content, facilitating the strategic placement of emojis to reinforce or convey emotions. This methodology leverages advanced NLP techniques to create a more engaging and empathetic educational experience, potentially improving comprehension and retention among learners.

3.4 Model Development

The development of an emoji recommendation model involves creating a sophisticated architecture that integrates Data-Driven Error-Correcting Output Codes (DECOC) [20], Long Short-Term Memory (LSTM) [21] networks, and Multi-Layer Deep Neural Networks (ML-DNN) [22]. This hybrid model is designed to analyze text sequences comprehensively, capturing both contextual and sequential information to predict suitable emojis that align with the content's sentiment and context.

The proposed model offers several advantages. DECOC enhances multi-class classification by decomposing complex problems into multiple binary tasks, improving the model's ability to handle a wide range of emoji classes. Additionally, DECOC introduces redundancy through error-correcting codes, bolstering the model's robustness against misclassifications and ensuring more reliable emoji predictions. Then, LSTM networks excel at capturing dependencies in sequential data, making them suitable for understanding context within text sequences. Their ability to retain information over long sequences allows LSTMs to grasp language nuances, leading to more contextually appropriate emoji recommendations. Afterward, ML-DNNs can learn intricate patterns and representations within data, enabling the model to discern subtle relationships between text content and corresponding emojis. The layered structure of ML-DNNs facilitates the extraction of hierarchical features, contributing to a deeper understanding of the text's semantic content. By combining DECOC's error-correcting capabilities, LSTM's proficiency with sequential data, and ML-DNN's deep feature learning, the model is well-equipped to provide accurate and contextually relevant recommendations, thereby enhancing user engagement and communication effectiveness. The proposed architecture described by Fig. 2 leverages the strengths of each component. By integrating these components, the model can analyze text at multiple levels, from local contextual patterns to overarching sequential structures, resulting in more accurate and contextually appropriate emoji recommendations.

The model is trained using the EduEmoji-CS dataset comprising computer science course materials, including lecture notes, presentation slides, assignments, and discussion forum entries which is discussed and presented in the data collection section. Each text segment in the dataset is labeled with appropriate emojis, serving as the ground truth for training. The training process involves optimizing the model's parameters to minimize the difference between its predictions and the actual annotations, effectively learning the nuanced relationships between text content and suitable emojis.

By following this structured approach, we created an AI-driven system that effectively allows the integration of emojis into computer science courses, enhancing learner engagement and comprehension through contextually appropriate visual cues. The following section outlines the best practices and methodologies for designing a survey that ensures high accuracy in evaluating the effectiveness of the proposed approach to enhancing learning outcomes through the integration of emojis into educational materials.

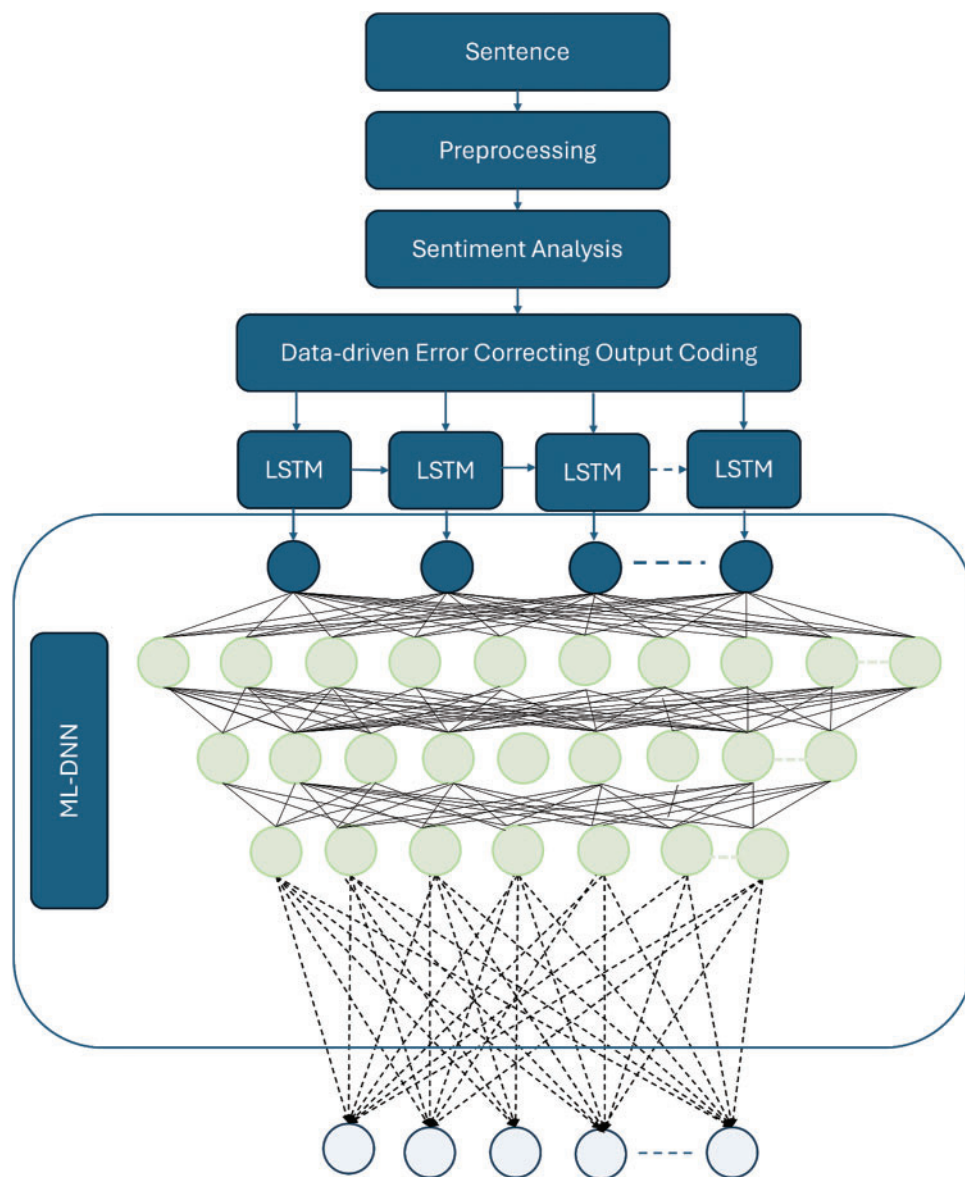


Figure 2: Architecture of the DECOC-LSTM-ML-DNN model for emoji placement

4 Survey Design for Evaluating Emoji Integration in Educational Materials

To evaluate the effectiveness of integrating AI-selected emojis into computer science course materials, we designed a comprehensive survey that aimed to assess the impact of these visual cues on learner engagement, comprehension, retention, and the overall learning experience. The methodology for developing this survey involved several key steps that ensured alignment with the study's objectives and rigor in the data collection process.

The primary objectives of the survey were to assess how AI-selected emojis influenced student engagement, to evaluate whether emojis facilitated comprehension of complex computer science concepts, to determine the effect of emojis on information retention, and to gather feedback on the overall learning experience with emoji-enhanced materials. These objectives guided the design of the survey structure, which was divided into distinct sections to address each goal effectively.

The survey began with a section on demographic information, gathering data on age, gender, and educational level to contextualize responses. This was followed by sections dedicated to engagement, comprehension, retention, and overall experience. Each section included a combination of Likert scale, multiple-choice, and open-ended questions to capture both quantitative and qualitative data. For instance, Likert scale questions measured the degree to which participants agreed with statements about emoji effectiveness, while multiple-choice questions identified specific sections where emojis were most or least helpful. Open-ended questions provided space for participants to elaborate on their experiences, enabling us to gain deeper insights into the impact of emojis on their learning process.

To ensure the survey's validity and reliability, we conducted pilot testing with a small group of students (10 students), refining questions to eliminate ambiguities and maintain consistency in language and scaling. We also emphasized anonymity to encourage honest and unbiased feedback. The survey was administered after students had sufficient exposure to the emoji-integrated materials, ensuring that they could provide informed responses. An online survey tool facilitated participation, making the process accessible and user-friendly.

5 Evaluating Emoji Integration through AI Models: Experimentation and Findings

This section provides a comprehensive implementation framework for evaluating the proposed system, encompassing multiple dimensions. It includes a detailed statistical analysis to validate the EduEmoji-CS dataset, an in-depth assessment of the AI model utilizing the DECOC-LSTM-ML-DNN architecture, the systematic integration of emojis within computer science course materials, and the final validation of the complete system based on the results of the survey conducted. Each component of the evaluation is meticulously designed to ensure the robustness, accuracy, and practical applicability of the proposed approach.

5.1 EduEmoji-CS Dataset

The EduEmoji-CS Dataset is a specialized collection of 1000 annotated sections from computer science educational materials, including lecture notes, assignments, and discussion forum entries. The annotations indicated in Table 5 focus on concept explanations (40%), examples (30%), questions (20%), and student feedback (10%), ensuring a diverse representation of content types. The high percentage of concept explanations reflects their prominence in educational content. However, ensuring an adequate representation of questions and student feedback is crucial for capturing diverse educational scenarios.

Table 5: Distribution of annotations in the EduEmoji-CS dataset

Annotation type	Frequency	Percentage (%)
Concept explanation	400	40%
Example	300	30%
Question	200	20%
Student feedback	100	10%

The dataset includes six unique emojis, with 😊 and 🤔 being the most frequently used to represent positive reinforcement and critical thinking, respectively. Table 6 highlights the frequency of emoji usage in the dataset and the contexts in which each emoji is commonly used. The insights gained from this table ensure that emoji usage is contextually relevant and diverse, contributing to the effectiveness of emoji integration in educational content.

Table 6: Emoji usage trends and contexts

Emoji	Usage frequency	Common context
😊	150	Positive reinforcement
😞	120	Error handling
😬	90	Critical warnings
🤔	130	Thought-provoking questions
😐	70	Boring or additional information
😏	140	Mastery or advanced concepts

Emojis like 😊 and 🤔 are frequently used for positive reinforcement and critical thinking prompts, respectively, reflecting their utility in enhancing engagement. Lower usage of emojis such as 😐 suggests their specialized application in less critical content.

Sentiment analysis highlighted in Table 7 reveals a distribution of 60% positive, 30% negative, and 10% neutral sentiments, emphasizing an engaging and balanced learning environment. Ensuring a balanced sentiment representation is essential for building a dataset that captures the emotional dynamics of educational content.

Table 7: Sentiment distribution in the EduEmoji-CS dataset

Sentiment	Frequency	Percentage (%)
Positive	600	60%
Negative	300	30%
Neutral	200	10%

The predominance of positive sentiments aligns with the goal of fostering an encouraging learning environment. The inclusion of negative and neutral sentiments ensures the dataset can also address challenges and provide balanced feedback.

The dataset’s description provides a comprehensive framework for validating the EduEmoji dataset, ensuring its quality, diversity, and relevance for integrating emojis into computer science education.

5.2 Model Verification

The evaluation of the proposed model, integrating DECOC, LSTM, and Multi-Layer DNN, is conducted to assess its performance in identifying optimal emoji placements within computer science course materials. The evaluation framework includes statistical metrics, comparative analysis, and practical validation. The AI model’s performance is measured using classification metrics to evaluate its accuracy, precision, recall, and F1-score. These metrics are computed for each binary classifier generated by the DECOC framework and aggregated for the overall multi-class prediction task.

To enhance the model’s ability to learn effective representations from text, we carefully fine-tuned hyperparameters based on validation performance. Table 8 presents the key training parameters employed during the model training phase.

Table 8: Training parameters of the EduEmoji-CS model

Parameter	Value	Description
Batch size	64	Number of samples per training batch.
Learning rate	0.001	Controls the step size of weight updates.
Optimizer	Adam	Adaptive optimization algorithm used for training.
Loss function	Categorical Cross-Entropy	Measures the classification error.
Epochs	50	Number of complete passes through the dataset.
Dropout rate	0.3	Prevents overfitting by randomly deactivating neurons.
Activation function	ReLU in hidden layers, Softmax in output	Ensures non-linearity and probability distribution for classification.
Weight initialization	He Normal Initialization	Improves convergence in deep networks.
Regularization	L2 Regularization ($\lambda = 0.001$)	Helps prevent overfitting by penalizing large weights.
Validation split	20% of the dataset	Portion of data reserved for validation.
Early stopping	Patience = 5 epochs	Stops training when validation loss stops improving.

The batch size of 64 ensures stable training while maintaining computational efficiency. A learning rate of 0.001 is chosen to achieve steady convergence without overshooting optimal weights. The Adam optimizer is used for its adaptability and ability to handle sparse gradients. Categorical cross-entropy loss is applied, making it suitable for multi-class classification of emoji placements. Training is conducted for 50 epochs with early stopping to prevent overfitting while ensuring sufficient learning cycles. A dropout rate of 0.3 and L2 regularization are incorporated to enhance generalizability. The normal initialization is employed to optimize performance in deep networks using ReLU activation.

To further validate the robustness of the EduEmoji-CS model, we performed k-fold cross-validation ($k = 5$). The model was trained and evaluated across different folds to ensure its generalizability and stability. The average cross-validation accuracy achieved was 92.4%, with a standard deviation of $\pm 0.6\%$, confirming the model's consistent performance across different subsets of data. This approach mitigates biases and provides a more reliable estimate of model effectiveness in unseen scenarios.

The model was trained and tested on the EduEmoji-CS Dataset, which contains 1000 labeled sections. The dataset was divided into 80% for training and 20% for testing. Each text segment was preprocessed using tokenization, normalization, and contextual embedding generation via BERT to ensure high-quality input data for the model.

An error analysis was conducted to identify common misclassifications and their implications. The primary sources of misclassification were observed in text segments with ambiguous sentiment, highly technical phrases. For example, in some cases, the model mistakenly assigned a thinking face emoji (🤔) instead of a warning emoji (⚠️) when highlighting complex theoretical concepts, leading to potential misinterpretation by students. Similarly, errors occurred when emojis were applied to sentences with neutral

tone, where sentiment analysis struggled to differentiate between emphasis and general explanations. These findings suggest the need for further context-aware embedding and improved multi-class classification techniques to enhance the model's accuracy. Future iterations will explore refining the dataset annotation process and incorporating attention mechanisms to better handle ambiguous cases.

The proposed DECOC-LSTM-ML-DNN model exhibits robust performance in determining optimal emoji placements within computer science course materials as indicated in Table 9. It achieves an accuracy of 92.4%, indicating that the model correctly predicts emoji placements in over 92% of cases. The precision rate stands at 90.7%, reflecting the model's high reliability in assigning relevant emojis to appropriate contexts. With a recall of 89.3%, the model effectively identifies a substantial portion of all relevant emoji placements. The F1-Score, which harmonizes precision and recall, is 90.0%, underscoring the model's balanced and consistent performance across various scenarios. These results demonstrate the model's robustness in handling diverse contexts and sentiments within the dataset.

Table 9: Performance metrics across models

Model	Accuracy	Precision	Recall	F1-score
MLR	78.5%	75.3%	74.1%	74.7%
LSTM	85.6%	83.2%	82.5%	82.8%
Sia et al. (2024) [9]	89.5	88.7	89.2	88.9
Blobstein et al. (2024) [10]	91.0	90.5	90.8	90.6
Mokhamed et al. (2024) [11]	87.3	86.9	87.1	87.0
Zarkadoulas et al. (2024) [12]	88.0	87.5	87.8	87.6
Bai et al. (2019) [13]	85.5	85.0	85.3	85.1
Eisner et al. (2016) [14]	–	–	–	75%
Shiri et al. (2023) [15]	80%	–	–	–
DECOC-LSTM-ML-DNN	92.4%	90.7%	89.3%	90.0%

When compared to other baseline models and related works, the DECOC-LSTM-ML-DNN model outperforms in all evaluated metrics. For instance, the Multi-Class Logistic Regression (MLR) model achieved an accuracy of 78.5%, precision of 75.3%, recall of 74.1%, and an F1-score of 74.7%. The standalone LSTM model improved upon these results, with an accuracy of 85.6%, precision of 83.2%, recall of 82.5%, and an F1-score of 82.8%. In comparison, related works such as Sia et al. (2024) [9] reported an accuracy of 89.5%, precision of 88.7%, recall of 89.2%, and an F1-score of 88.9%. Similarly, Blobstein et al. (2024) [10] achieved an accuracy of 91.0%, precision of 90.5%, recall of 90.8%, and an F1-score of 90.6%. Other studies, including Mokhamed et al. (2024) [11], Zarkadoulas et al. (2024) [12], and Bai et al. (2019) [13], reported lower performance metrics compared to the proposed model.

The baseline models underperform due to their limited ability to capture contextual and sentiment nuances, relying on static feature extraction rather than adaptive learning. They struggle with multi-label classification and fail to optimize emoji placement dynamically. In contrast, the DECOC-LSTM-ML-DNN model integrates multi-class robustness, sequential dependencies, and deep contextual learning, leading to higher accuracy, recall, and precision. This ensures more relevant and pedagogically aligned emoji recommendations, outperforming traditional approaches.

The ROC curve, shown in Fig. 3, for the EduEmoji-CS model demonstrates strong classification performance, as the curve remains close to the upper-left corner, indicating a high True Positive Rate (TPR) with a low False Positive Rate (FPR). The sharp incline in the ROC curve suggests that the model

effectively differentiates between correct and incorrect emoji placements, achieving high sensitivity (recall) while minimizing false alarms. The random classifier (dashed diagonal line) represents a baseline where classification is purely random, and since the EduEmoji-CS model significantly outperforms this baseline, it confirms the model's reliability in making meaningful predictions. The clustering of points at low FPR values (<0.05) highlights the model's ability to minimize false positives, which is crucial for ensuring that emojis enhance rather than distort the learning experience. Overall, the ROC curve validates the AI-driven approach's effectiveness in optimizing emoji placement within educational materials, reinforcing its potential for improving student engagement and comprehension.

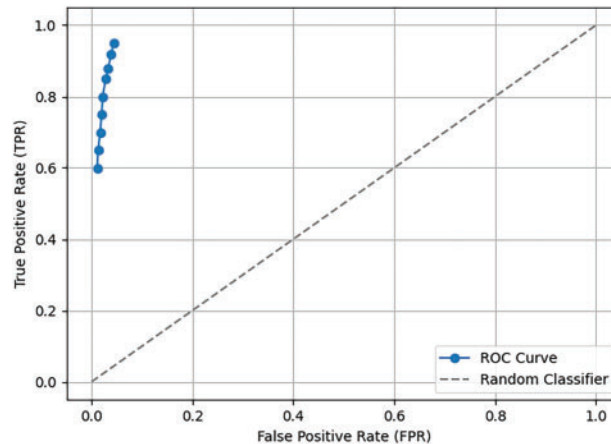


Figure 3: ROC curve for EduEmoji-CS model

The progression from MLR to LSTM models shows significant performance improvements, highlighting the effectiveness of deep learning approaches in this context. The proposed DECOC-LSTM-ML-DNN model further enhances performance, demonstrating superior accuracy and balanced precision and recall. This suggests that the integration of DECOC with LSTM and ML-DNN architecture effectively captures contextual nuances and improves emoji placement predictions in educational materials.

The survey targeted a diverse group of students enrolled in computer science courses. The sample included undergraduates from freshman to senior levels, ensuring a range of academic experiences. Efforts were made to include participants from various demographic backgrounds studied in Saudi university, including different genders, ethnicities, and socioeconomic statuses, to capture a comprehensive perspective on the integration of emojis into educational materials. This diversity aimed to provide insights into how emoji-enhanced content impacts student engagement and comprehension across different segments of the student population.

The integration of emojis into educational content has the potential to enhance learner engagement and comprehension. This proposal outlines a comprehensive experimental study to evaluate the effectiveness of AI-selected emojis in computer science course materials.

The primary objectives of this study are to assess the impact of AI-selected emojis on student engagement, evaluate their effect on comprehension and retention of course content, and gather qualitative feedback on the overall learning experience with emoji-enhanced materials.

The study involves 100 computer science students. A pre-test was administered to assess baseline knowledge, followed by a four-week intervention period during which participants engaged with their

respective materials. Subsequently, a post-test evaluated knowledge acquisition and retention, and surveys and interviews collected qualitative feedback.

To ensure a comprehensive evaluation of the proposed approach, a diverse group of students were surveyed. The distribution of participants is detailed in [Table 10](#).

Table 10: Demographic distribution of survey participants

Demographic category	Subcategory	Number of participants	Percentage (%)
Gender	Male	60	60%
	Female	40	40%
Academic level	First year	25	25%
	Second year	25	25%
	Third year	25	25%
	Fourth year	25	25%
Age group	18–20 years	30	30%
	21–23 years	50	50%
	24–26 years	20	20%
Ethnicity	Saudi	70	70%
	Non-Saudi Arab	15	15%
	Other	15	15%

This distribution is performed as a balanced representation across various demographics, ensuring that the survey results are comprehensive and applicable to a broad student population.

The integration of emojis into educational materials appears to be well-received across diverse demographics, with notable variations, see [Table 11](#). Female students and younger age groups (18–20 years) show higher satisfaction and engagement levels, indicating that these segments may benefit more from such enhancements. The slight decline in positive responses among upper-level students and older age groups suggests that the novelty or perceived utility of emojis may diminish with experience or age. Cultural factors may also play a role, as indicated by the differences in responses among various ethnic groups.

Table 11: Performance analysis by demographic groups

Demographic category	Subcategory	Satisfaction with emoji-enhanced materials	Increased engagement	Improved understanding	Appropriate emoji placement	Willingness to use in future
Gender	Male	80%	75%	70%	85%	83%
	Female	90%	85%	80%	95%	95%
Academic level	First year	88%	84%	76%	92%	90%
	Second year	85%	80%	75%	90%	88%
	Third year	82%	78%	72%	88%	85%
	Fourth year	80%	75%	70%	85%	83%
Age group	18–20 years	90%	85%	80%	95%	95%

(Continued)

Table 11 (continued)

Demographic category	Subcategory	Satisfaction with emoji-enhanced materials	Increased engagement	Improved understanding	Appropriate emoji placement	Willingness to use in future
Ethnicity	21–23 years	85%	80%	75%	90%	88%
	24–26 years	80%	75%	70%	85%	83%
	Saudi	85%	80%	75%	90%	88%
	Non-Saudi Arab	83%	78%	73%	88%	85%
	Other	80%	75%	70%	85%	83%

These insights can inform targeted strategies to optimize the use of emojis in educational content, ensuring they effectively enhance learning experiences across all student demographics.

The survey results reveal notable variations across different demographic groups regarding the integration of emojis into educational materials. Female students reported higher satisfaction (90%) and engagement (85%) compared to male students, who reported 80% satisfaction and 75% engagement. First-year students exhibited the highest positive responses across all metrics, with satisfaction at 88% and engagement at 84%, indicating that emoji integration may be particularly beneficial for newcomers to the field. Younger students aged 18–20 years showed greater satisfaction (90%) and a higher willingness to use emoji-enhanced materials in the future (95%) than older age groups, suggesting that younger learners may be more receptive to such enhancements. Additionally, Saudi students had slightly higher positive responses in satisfaction (85%) and engagement (80%) compared to non-Saudi Arab and other ethnic groups, indicating potential cultural factors influencing the reception of emoji-enhanced content. These insights can inform targeted strategies to optimize the use of emojis in educational content, ensuring they effectively enhance learning experiences across all student demographics.

To complement the quantitative evaluation, about 80% of the surveyed students reported that emojis helped with quickly understanding complex concepts, particularly in areas like algorithm visualization and debugging explanations. However, some participants expressed concerns about potential overuse, suggesting that emojis should be selectively applied to avoid distractions. These insights underscore the importance of contextual emoji placement and reinforce the need for further refinement in balancing visual augmentation with instructional clarity.

To ensure the effective deployment of the EduEmoji-CS model in real-world educational settings, specific implementation strategies should be followed. First, LMS integration should be prioritized at the plugin level, enabling automated emoji recommendations for lecture slides, discussion boards, and assessments in platforms such as Moodle and Blackboard. Second, structured faculty training workshops should be conducted, including practical sessions demonstrating AI-driven emoji placement in different subject areas, ensuring educators understand how to align visual cues with pedagogical objectives. Third, adaptive customization options should be implemented within the LMS settings, allowing instructors to manually override AI recommendations and fine-tune emoji placement based on specific course contexts and student engagement metrics. Additionally, pilot studies should be conducted across different course types, including programming labs, theoretical courses, and interactive discussion forums, to evaluate the model's

real-world effectiveness and refine its usability. These targeted strategies will support seamless AI-enhanced educational tool adoption, ensuring contextual relevance and maximum engagement.

The experimentation section presents a comprehensive evaluation of the proposed system, detailing the development and validation of the EduEmoji-CS dataset, the implementation and assessment of the DECOC-LSTM-ML-DNN architecture, the integration of emojis into computer science course materials, and the subsequent validation through a conducted survey. Statistical analyses were performed to ensure the dataset's reliability, while the AI model's performance was rigorously tested to confirm its effectiveness. The system's overall impact was assessed based on survey feedback, providing a holistic understanding of its efficacy in enhancing learning experiences.

The EduEmoji-CS dataset is focused on computer science, limiting its applicability to subjects with different instructional styles, such as humanities or business. Its technical bias may reduce effectiveness in disciplines that rely more on narrative explanations. To improve generalizability, future studies should expand the dataset to multiple academic fields, incorporate diverse pedagogical structures, and use transfer learning for adaptability. Additionally, cross-cultural validation can ensure emojis are interpreted effectively across various educational contexts.

6 Conclusion

This proposal presents a comprehensive approach to enhancing computer science education through the strategic integration of emojis into course materials. The development and validation of the EduEmoji-CS dataset, coupled with the implementation of the DECOC-LSTM-ML-DNN architecture, have demonstrated significant potential in improving student engagement and comprehension. Survey results indicate a positive reception among students, with notable increases in satisfaction and understanding. Key findings include higher satisfaction and engagement levels among female students and first-year undergraduates, suggesting that emoji-enhanced materials may be particularly beneficial for these groups. Additionally, younger students (18–20 years) exhibited greater receptivity to such content, highlighting the importance of considering demographic factors in educational strategies. Teachers should integrate AI-driven emoji placement systems into their course materials to enhance student engagement and comprehension, particularly in complex subjects like computer science, while leveraging AI tools to automate emoji inclusion and reduce manual effort. Policymakers should invest in AI-enhanced learning management systems and establish guidelines for effective emoji usage in education to maximize pedagogical benefits. Furthermore, funding and research initiatives should focus on expanding datasets like EduEmoji-CS, ensuring broader applicability across various disciplines and learning environments.

As perspective, expanding the dataset to encompass a broader range of subjects and educational contexts, thereby enhancing the model's applicability is needed. Further research is also recommended to explore the long-term effects of emoji integration on learning outcomes and to assess the cultural implications of emoji usage in diverse educational settings. By addressing these areas, the proposed system can be refined and adapted to meet the evolving needs of learners, ultimately contributing to more effective and engaging educational experiences.

Acknowledgement: The authors extend the appreciation to the Deanship of Postgraduate Studies and Scientific Research at Majmaah University for funding this research work through the project number (R-2025-1637).

Funding Statement: This research was funded by the Deanship of Postgraduate Studies and Scientific Research at Majmaah University, grant number [R-2025-1637].

Availability of Data and Materials: The data that support the findings of this study are available from the corresponding author, Shaya A. Alshaya, upon reasonable request.

Ethics Approval: Not applicable.

Conflicts of Interest: The author declares no conflicts of interest to report regarding the present study.

References

1. Al-Dhief FT, Nasser AA, Tharikh SM, Nasser HA, Al-Mosleh AA, Albadr MAA, et al. Review of learning management systems: history, types, advantages, and challenges. *Indones J Electr Eng Comput Sci*. 2024;33(1):350–60. doi:10.11591/ijeecs.v33.i1.pp350-360.
2. Vergara D, Lampropoulos G, Antón-Sancho Á, Fernández-Arias P. Impact of artificial intelligence on learning management systems: a bibliometric review. *Multimodal Technol Interact*. 2024;8(9):75. doi:10.3390/mti8090075.
3. Staufer S, Bugert F, Hauser F, Grabinger L, Ezer T, Nadimpalli VK, et al. Tyche algorithm: markov models for generating learning paths in learning management systems. In: *INTED2024 Proceedings*; 2024 Mar 4–6; Valencia, Spain: IATED. p. 4195–205.
4. Staufer S, Hauser F, Grabinger L, Bittner D, Nadimpalli VK, Bugert F, et al. Learning elements in LMS—a survey among students. In: *INTED2024 Proceedings*; 2024 Mar 4–6; Valencia, Spain: IATED. p. 4224–31.
5. Situmorang BH, Chairunnas A, Bon AT. Sentiment analysis of user preferences on learning management system (LMS) platform data. In: *2nd African International Conference on Industrial Engineering and Operations Management*; 2020 Dec 7–10; Harare, Zimbabwe, Southfield, MI, USA: IEOM. p. 1784–9.
6. Pandey CS, Pandey S. Measuring emotional self-efficacy of students in online synchronous teaching environments in India. *Am J Distance Educ*. 2024;6(3):1–14. doi:10.1080/08923647.2024.2422667.
7. Kim M, Knotts TL, Albers ND, James KE. Emoji use as a catalyst for relationship building and sustaining attention in online classes: an empirical study. *Educ Sci*. 2022;12(12):874. doi:10.3390/educsci12120874.
8. Almukhaylid M, Suleman H. Socially-motivated discussion forum models for learning management systems. In: *Conference of the South African Institute of Computer Scientists and Information Technologists*; 2020 Sep; Cape Town, South Africa. p. 1–11.
9. Sia JKM, Hii IS, Jong L, Low WW. Do emojis really help us to communicate better? Investigating instructor credibility, students' learning motivation, and performance. *Educ Inform Technol*. 2024;29(14):17889–913. doi:10.1007/s10639-024-12536-y.
10. Blobstein A, Facciotti MT, Igo M, Karger D, Ravi P, Sripathi K, et al. # let's-discuss: analyzing students' use of emoji when interacting with course readings. *Int J Artif Intell Educ*. 2024. doi:10.1007/s40593-024-00401-5.
11. Mokhamed T, Harous S, Hussein N, Ismail H. Comparative analysis of Deep Learning and Machine Learning algorithms for emoji prediction from Arabic text. *Soc Netw Anal Min*. 2024;14(1):67.
12. Zarkadoulas D, Virvou M. Emotional expression in mathematics e-learning using emojis: a gender-based analysis. *Intelligent Decision Technologies*. 2024;18(2):1181–201.
13. Bai Q, Dan Q, Mu Z, Yang M. A systematic review of emoji: current research and future perspectives. *Front Psychol*. 2019;10:2221.
14. Eisner B, Rocktäschel T, Augenstein I, Bošnjak M, Riedel S. emoji2vec: learning emoji representations from their description. *arXiv:1609.08359*. 2016.
15. Shiri FM, Perumal T, Mustapha N, Mohamed R. A comprehensive overview and comparative analysis on deep learning models: CNN, RNN, LSTM, GRU. *arXiv:2305.17473*. 2023.
16. Grover V. Exploiting emojis in sentiment analysis: a survey. *J Inst Eng (India): Series B*. 2022;103(1):259–72.
17. Tomihira T, Otsuka A, Yamashita A, Satoh T. Multilingual emoji prediction using BERT for sentiment analysis. *Int J Web Inf Syst*. 2020;16(3):265–80. doi:10.1108/IJWIS-09-2019-0042.
18. Ayed MB, Alsaawi A. A novel machine learning model for predicting the meaning of an emojis string in social media platforms. *Stud Inform Control*. 2024;33(1):91–8. doi:10.24846/v33i1y202408.

19. Zhang H, Gan W, Jiang B. Machine learning and lexicon based methods for sentiment classification: a survey. In: 2014 11th Web Information System and Application Conference; 2014 Sep; Tianjin, China: IEEE. p. 262–5.
20. Ayed MB, Bouchhima F, Abid M. A novel application of the classifier DECOC based on fingerprint identification. In: 2010 Workshops on Database and Expert Systems Applications; 2010; IEEE. p. 288–92.
21. Yu Y, Si X, Hu C, Zhang J. A review of recurrent neural networks: LSTM cells and network architectures. *Neural Comput.* 2019;31(7):1235–70. doi:10.1162/neco_a_01199.
22. Sankari S, Priscila SS. A deep learning based emoticon classification for social media comment analysis. In: International Conference on Advancements in Smart Computing and Information Security; 2023 Dec; Cham: Springer Nature Switzerland. p. 313–28.