

Leveraging Deep Learning to Optimize AI Self-Efficacy Development through Tailored Prompt Engineering Interventions

N. Saraswathi¹, Dr. K. Gurusamy^{1*}, Dr. N. Jenipha², S. Banupriya³, Dr. N. Chitra⁴

Research Scholar, Department of English, Kalasalingam Academy of Research and Education, Tamil Nadu, India. 626126.

Associate Professor, Department of English, Kalasalingam Academy of Research and Education, Tamil Nadu, India. 626126.

Assistant Professor (Sl.Gr.), Department of English, University College of Engineering
Tamil Nadu, India. 626126.

Assistant Professor, Department of English, Kalasalingam Academy of Research and Education, Tamil Nadu, India. 626126.

Assistant Professor (SL.GR.), Department of English, UCE, BIT Campus, Anna University Tamil Nadu, India. 626126.

Corresponding Author: Dr. K. Gurusamy,

HOW TO CITE:

N. Saraswathi, K. Gurusamy, N. Jenipha, S. Banupriya, N. Chitra (2026). Leveraging Deep Learning to Optimize AI Self-Efficacy Development through Tailored Prompt Engineering Interventions International Journal of Special Education, 41(22s), 1060 - 1069

COPYRIGHT STATEMENT:

Copyright: © 2026 Authors.
Open access publication under the terms and conditions the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

ABSTRACT:

Artificial Intelligence (AI) self-efficacy is an essential component in the transformation of education, especially in improving the learning outcomes among students, as AI can do. Nevertheless, existing systems usually do not offer individualized learning opportunities, which restricts the potential of students. The proposed study will present a Position Attention -based Long Short-Term Memory (PA-LSTM) model, which identifies the level of AI self-efficacy in students based on the processing of real-time interaction information. The model is a combination of temporal modeling and position-based attention, which considers both behavioral and contextual aspects of interactions between students and AI. Such data preprocessing methods as cleaning, normalization, and text tokenization guarantee the provision of reliable and stable input to the model. The extraction of the features incorporates the student prompts and behavioral indicators including response time, edits and self-efficacy ratings. The proposed model has always shown better performance, with the accuracy of 97.5% to 99.5%. These findings underscore the strength of the model to give individualized feedbacks and effectively categorize learning outcomes, therefore, improving AI self-efficacy building in classrooms.

Keywords: Real-Time Data, Position Attention Mechanism, Long Short-Term Memory, Personalized Learning

1. Introduction

AI has quickly swept through various industries, such as education, healthcare, and business, making it possible to automate, offer personal experiences, and solve problems on a new level [1]. In the educational setting, students should not simply be knowledgeable of theoretical concepts of AI, but should also learn to work with the systems of AI. One of the most important competencies in this respect is the AI self-efficacy or the belief in the ability to use AI tools to attain the desired outcomes [2]. By increasing the capacity of students to become effective prompts, they can use AI products to their advantage, engage more, and enhance their knowledge and practical skills of AI, finally increasing their learning outcomes and confidence [3].

There are a number of constraints on the development of AI self-efficacy among learners. First, AI systems can be confusing and complicated, and beginners will be scared and unable to experiment [4]. Second, traditional instructional patterns tend to be homogeneous in terms of teaching approaches and do not take into consideration the differences in cognitive styles, the level of previous knowledge, or engagement [5]. Third, students do not have access to prompt, personal feedback on their immediate engineering attempts and may make the same errors repeatedly, and lose motivation. All these collectively lower the confidence of the learners and slow the process of learning skills, thus, there is need to implement adaptive, responsive interventions which reshape the learning tasks and instructions in real time in line with the capabilities and progression of each student [6].

Some techniques have been suggested to enhance AI learning and elicit engineering skills. Classical workshops and tutorials are structured, and they offer little customization [7]. Adaptive guidance systems are combined with Intelligent Tutoring Systems (ITS) and reinforcement learning-based interfaces, while student progress is tracked using predictive analytics based on Machine Learning (ML) models, such as Support Vector Machines

(SVM), Random Forests (RF), and Decision Trees (DT) [8]. Also, the strategies ignore semantic richness of student prompts and thus, it is hard to effectively measure skill development expressively. Moreover, the majority of the systems provide fixed feedback, which cannot adjust to the changing needs of the student in real time, thus not providing enough motivation and encouragement in the learning process [9].

An innovative framework suggested in the proposed research solves the shortcomings of the current systems. This framework can enable the improved classification of AI self-efficacy by combining semantic analysis with a PA-LSTM network, processing sequential interaction, and semantic depth of student prompts. The PA-LSTM model can provide real-time and personalized feedback with dynamic suggestions of personalized prompts and hints according to individual progress. The originality of the research is that it combines semantic understanding with temporal modeling and position-based attention, which seems to be a more receptive and context-sensitive way of developing AI self-efficacy.

2. Literature Survey

Within recent years, the problem of AI being introduced into the education sphere has been discussed in the context of several disciplines that have aimed at improving student engagement, learning experience, and self-efficacy. Research on Occupational Health and Safety (OHS) education has emphasized the need to equip engineering students to handle safety standards well. As opposed to conventional methods of educating such as textbooks, which may not be enough to respond to short-term and specific questions, AI chatbots have proven to be interactive tools that can now give real-time responses. A quasi-experimental study that compared the custom chatbot with the traditional methods had better learning performance, self-efficacy, and technology acceptance [10].

Nevertheless, inquiry-based learning, which was acclaimed as the method that facilitates deeper understanding, usually encounters problems

caused by the lack of curiosity and self-efficacy of students. To address this, the Automated Inquiry Prompt System (AIPS) was brought into being, combining ChatGPT with vision technology. The system keeps track of students' actions automatically and gives them prompts that are contextually appropriate, thus activating students' imagination, reasoning, and involvement [11].

In yet another study, the huge potential of Language Models (LLMs) to help the teaching of programming in Grade 7 was investigated. After six weeks of teaching via LLM, the control group of 103 students had a significant increase in programming performance, cognitive interest, and self-efficacy. The researchers also highlighted the importance of human-machine interaction and self-directed learning in real-time, which generated exploratory behaviors. It was concluded that LLMs can give rise to competence and autonomy in digital learning environments [12].

With the development of AI technology, personalization gains a greater role in education, particularly in knowledge areas such as engineering. There was evidence that multimodal methods combined with the use of Game-Based Learning Environments (GBLEs) can be used to produce customized learning experiences. Self-Regulated Learning (SRL) needs of students can be satisfied with the help of AI due to its ability to analyze diverse information, including log files, eye movements, facial expressions, and physiological data. This individual process is assistive to SRL processes and assists educators in comprehending the learning requirements of the students and developing more efficient education plans [13].

Also, the emergence of Large-Language Model-based AI Services (LLMAIs) has introduced new issues regarding the interpretation of the self-efficacy of users and their particular demands when using these technologies. The current tools were considered inadequate to measure the

dimensions, which led to the creation of the Questionnaire of Self-efficacy and Needs in using Core Features of LLMAIs (Q-SNELL). The designed tool based on the reviews of the professionals and surveys of the users and set to determine how well the LLMAIs can help with the improvement of the learning outcomes and the future developments [14].

Lastly, a literature review on AI literacy in college students came up with the AI Literacy Conceptual Model (AICM), which AI literacy, ethical implications, and affective issues such as motivation and self-efficacy to proposed in a multidimensional model. A self-report questionnaire revealed the importance of motivation in promoting AI literacy and self-efficacy and self-efficacy as a mediator between the two. It also indicated gendered differences in the dynamics of learning to the male students revealed higher mediation effects, and previous barriers to learning included the lack of AI experience and curricular support. This paper highlighted the need to attend to these issues to establish the enhancement of AI literacy [15].

3. Proposed Methodology

Figure 1 describes the AI self-efficacy classification methodology workflow of a Position Attention-Based LSTM model. The task starts with Data Collection, in which real-time data of student-AI interaction is collected. Then, Data Preprocessing guarantees the quality of data with the usage of cleaning, missing values, normalization and text tokenization. The data are refined and then Feature Extraction is done using the synthesized semantic and behavioral features to constitute meaningful input vectors. Lastly, the Classification Using Position Attention-Based LSTM step applies these sequences to detect temporal and contextual patterns, and classifies the learning outcomes as either High, Medium, or Low, and thus, improves individualized knowledge on student self-efficacy formation.

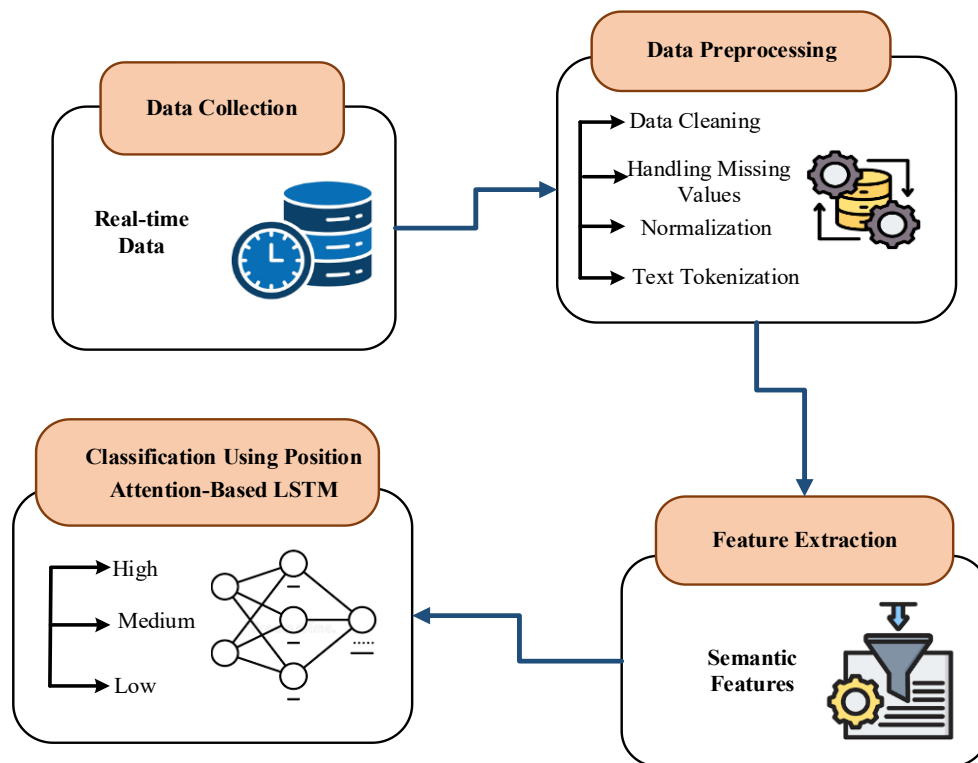


Figure 1. Overall proposed methodology

3.1. Data Collection

The input data were collected in real-time from prompt engineering workshops where students interacted with AI systems to create and calibrate prompts. The dataset included several aspects of interaction such as student-written prompts, responses generated by AI, and system interaction records such as time taken for each activity, the number of edits, and retry attempts. Micro self-efficacy survey scores were also captured after each activity to device the learner's confidence and perceived levels of competence. This close real-time monitoring guaranteed the collection of both cognitive and behavioral markers, providing a robust foundation for follow-on feature extraction and model training.

3.2. Data Preprocessing

Data preprocessing is used to deal with inconsistencies in raw interaction data, including typographical errors, missing fields and inconsistent scales. It converts input data into a standard format of machine learning models. The major preprocessing activities are data cleaning, processing missing values, normalization and the process of text tokenization. Data cleaning

consists of removing invalid records like missing, broken, or duplicated data that can affect the consistency of the dataset. The treatment of missing values is through mean imputation, which means that the missing values are replaced by the arithmetic mean of the available data for that feature. To normalize the features, min-max scaling is used to scale the features such as task duration, attempts, and engagement scores to the range $[0, 1]$ to carry out the normalization process. The normalization is shown by Equation (1):

$$x'_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \quad (1)$$

where, x_i is the normalized unique feature, $x_{\min}$ and $x_{\max}$ represent the minimum and maximum values. Text tokenization is the process that separates the text prompts and responses into units of words, thus allowing the downstream models to have a semantic representation.

3.3. Feature Extraction using Semantic Features

From the preprocessed data, each student interaction at time t is coded by a composite

feature vector with a combination of semantic and behavioral data, and is provided by Equation (2):

$$f_t = [v_t, E_t, T_t, A_t, S_t] \in \mathbb{R}^{d+4} \quad (2)$$

Here, v_t is used to refer to the prompt's semantic embedding, while the metrics of behavior measured by E_t, T_t, A_t, S_t account for the number of edits, response time, attempts, and self-efficacy scores. To visualize the temporal patterns of learning in students, these feature vectors have to be put in a sequential manner which will then constitute the LSTM network's input sequence as shown by Equation (3):

$$F = [f_1, f_2, \dots, f_N] \quad (3)$$

where, N is the number of interactions of a student. This temporal encoding enables the model to encode semantic meaning in questions as well as the temporal dynamics of learner behavior and hence provide a vigorous input to

downstream classification with the Position Attention-based LSTM.

3.4. Classification using Position Attention-Based LSTM

Figure 2 illustrates the model structure that PA-LSTM follows in learning outcomes classification. The first one is the Input Layer, where the interactive data of the students (prompts, answers and behavioral features) are typed. The data are subsequently transferred to three LSTM Layers (LSTM Layer-1, Layer-2, and Layer-3), which can learn the temporal patterns of the behavior of the learner by successively processing information. After the processing, the Position Attention Mechanism is performed, assigning dissimilar time steps in the sequence variable degrees of importance. Ultimately, the information is passed on using a Fully Connected Layer with SoftMax that provides a classification of learning outcomes as High, Medium and Low.

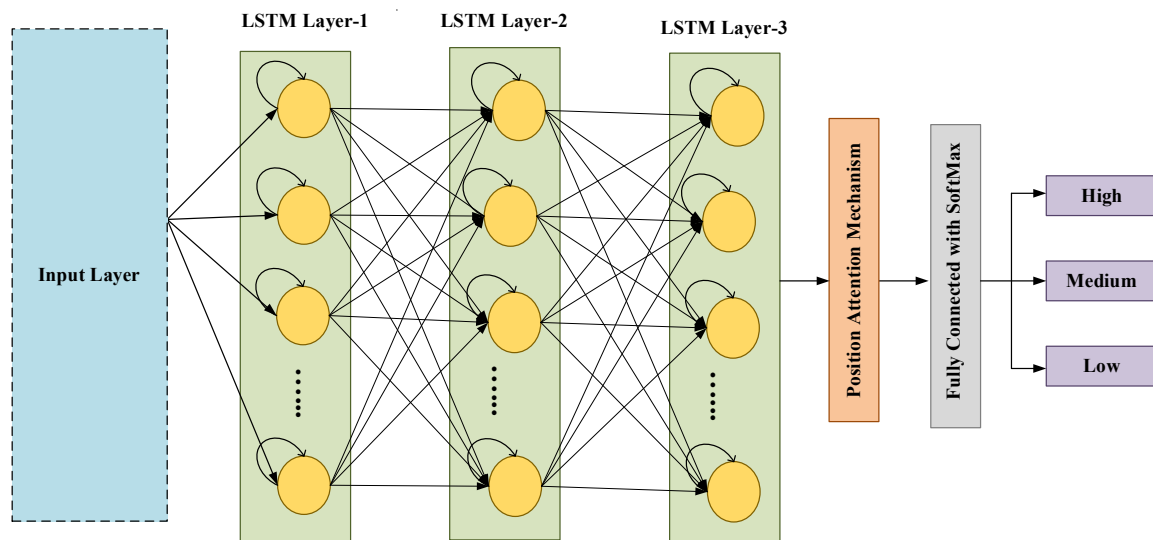


Figure 2. Architecture of PA-LSTM

The PA-LSTM model is a type of sequential data processing system that takes student interactions as input to classify learning outcomes and self-efficacy improvement. The LSTM cell calculates at the time step t the input gate i_t , the forget gate f_t , and the output gate o_t . The gates control the amount of current input and previous state information that should be kept or changed. The input gate is calculated as Equation (4):

$$i_t = \sigma(W_i f_t + U_i h_{t-1} + b_i) \quad (4)$$

where, W_i and U_i represent the weight matrices, f_t denotes the present input, h_{t-1} signifies the past hidden state, and b_i corresponds to the bias. The forget gate determines what portions of the preceding cell state c_{t-1} are to be reserved in Equation (5):

$$f_t = \sigma(W_f f_t + U_f h_{t-1} + b_f) \quad (5)$$

The output gate determines what proportion of the current cell state should be exposed as the

output of the LSTM at this time step as shown by Equation (6):

$$o_t = \sigma(W_o f_t + U_o h_{t-1} + b_o) \quad (6)$$

The LSTM's memory, the cell state c_t , gets its new value from the former cell state controlled by the forget gate and the new candidate cell state $\tilde{c}_t$ ruled by the input gate is define by Equation (7):

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \quad (7)$$

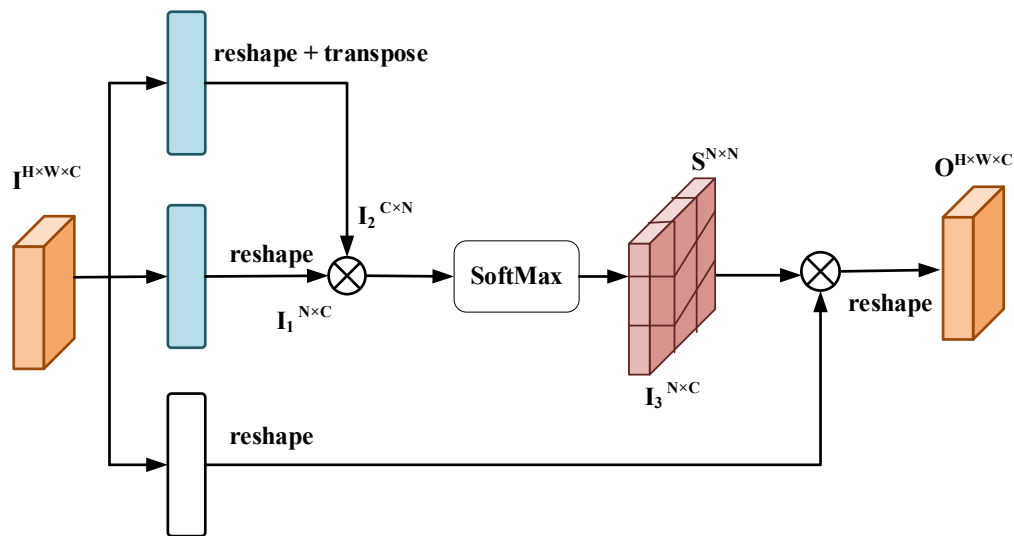


Figure 3. Position Attention Module

The Position Attention Module for calculating attention weights in sequential data is illustrated in Figure 3. The procedure starts with the resizing of the input matrix I_1 with size $N \times C$ and then its transposition leading to the matrix I_2 being formed, having dimensions $C \times N$. Next, element-wise multiplication with the query matrix H (size $H \times W \times C$) is performed. The output is subsequently subjected to the SoftMax function for scoring normalization that results in the generation of an attention matrix S , of size $N \times N$, which is indeed a probability distribution over the sequence. The computation of the attention weights α_t is as follows by Equation (9):

$$\alpha_t = \frac{\exp(W_a h_t)}{\sum_{k=1}^N \exp(W_a h_k)} \quad (9)$$

Here, W_a is the matrix of attention weights and h_t denotes the hidden state at time t . The softmax function is applied by the attention mechanism to the weights and, the model to

The hidden state h_t , the output for the current time step, is determined by Eq (8):

$$h_t = o_t \odot \tanh(c_t) \quad (8)$$

The weight α_t is assigned to each hidden state h_t based on the processing done through the LSTM cells by the PA Mechanism, which reflects the role that each time step is classifying the learning outcome.

concentrate on the most important interactions within the sequence. The specification of the attended representation h_{att} is done through the weighted sum of all hidden states as shown by Equation (10):

$$h_{att} = \sum_{t=1}^N \alpha_t h_t \quad (10)$$

The attended representation serves to bring forth the pertinent interactions, reducing the sequence that is ready for classification. Subsequently, the attended representation is conveyed through a SoftMax layer to determine the final classification output. The SoftMax function changes the attended representation as shown in Equation (11):

$$y = \text{softmax}(W_s h_{att} + b_s) \quad (11)$$

Here, W_s indicates the weight matrix corresponding to the classification layer, b_s indicates the bias, y indicates the output probabilities. The SoftMax function guarantees

the summation of the classified probabilities to 1; the model can categorize the student's learning outcome as High, Medium, or Low.

4. Results and Discussion

The experimental findings show that the proposed PA-LSTM model is effective in making correct classification of the learning outcomes and the level of self-efficacy of students.

Evaluation against several baseline Deep Learning (DL) models proves that the proposed model is more successful than them. The PA-LSTM is an efficient model that reflects temporal and semantic dependence of interaction between students and AI, which demonstrates its strength and impartiality to learn individual learning dynamics in educational settings.

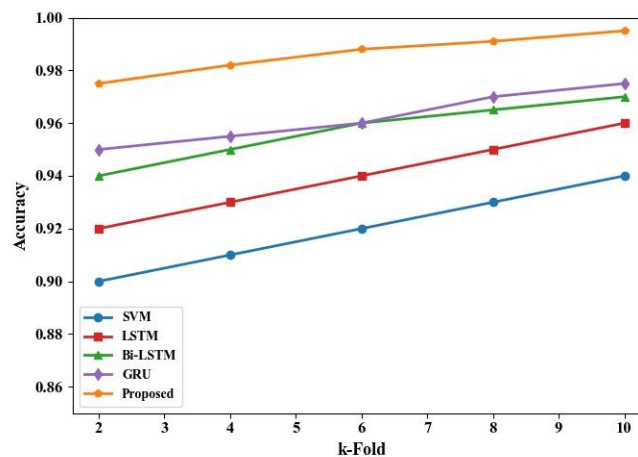


Figure 4. Accuracy

Figure 4 shows the accuracy of five models (SVM, LSTM, Bi-LSTM, GRU, and Proposed) in 10 k-Folds. SVM has an accuracy of .90 at the start of k-Fold and grows to 0.94 at k-Fold 10, LSTM performs at 0.92 to 0.96; Bi-LSTM is 0.94 to 0.97 and GRU goes from .95 to 0.975. The

Proposed model has the best accuracy going from 0.975 to the end point of k-Fold at 0.995. The Proposed model consistently has the best performance across k-Folds, especially between k-Fold 6 to the end, when it achieves near perfect accuracy.

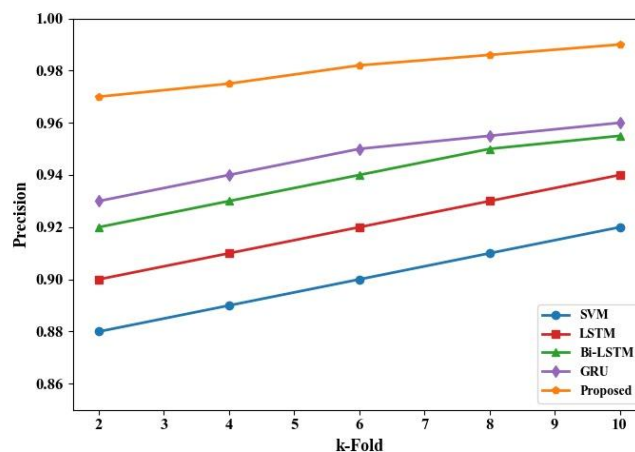


Figure 5. Precision

Figure 5 shows the precision of five models on 10 k-Folds. The initial precision of SVM is 0.88 and rises to 0.92 at k-Fold 10. LSTM begins at 0.90 and rises to 0.94. Bi-LSTM starts at 0.92 and

reaches 0.955. GRU starts at 0.93 and ends at 0.96. The Proposed model gives the best results with starting 0.97 and continuing to rise up to 0.99 at K-Fold 10. The Proposed model

continues its performance progressively better than the others especially starting with the 6 k-

Fold, where its performance steadily rises starting with almost ideal precision.

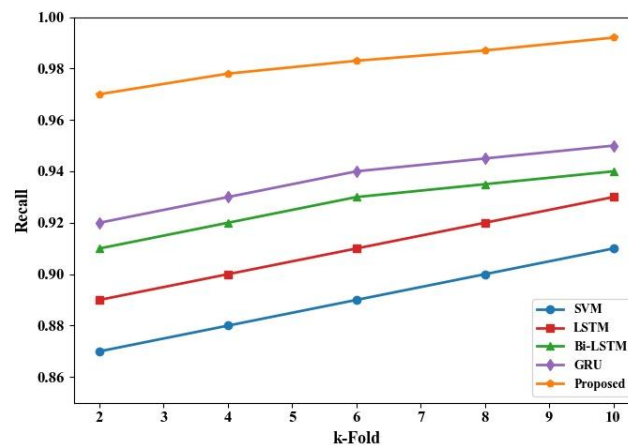


Figure 6. Recall

Figure 6 presents the recall of five models (SVM, LSTM, Bi-LSTM, GRU and Proposed) in 10 k-Folds. The initial SVM recall is 0.87 and it rises to 0.91 at k-Fold 10. LSTM begins at 0.89 and reaches 0.93. Bi-LSTM begins with 0.91 and goes up to 0.94. GRU begins at 0.92 and ends at 0.95. The Proposed model has the highest

performance with initial results of 0.97 and increasing to 0.992 in the 10th k-Fold. The Proposed model is always performing better than the rest, and the performance is steadily increasing, particularly, after the k-fold 6, the model achieves almost perfect recall.

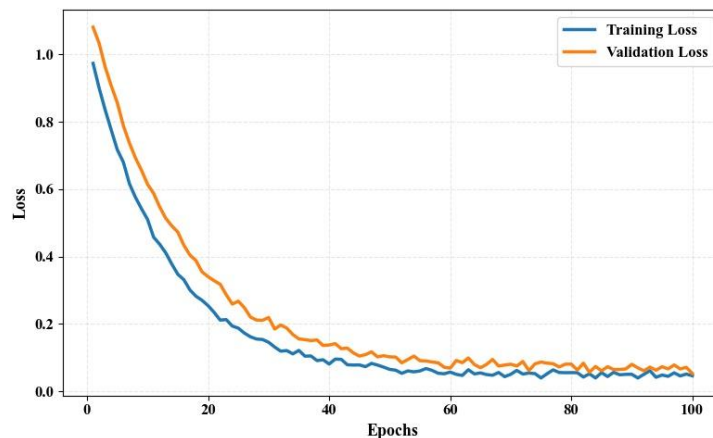


Figure 7. Training and validation loss

Figure 7 shows the values of the losses that occur in the training and validation process after 100 epochs. First, the training and the validation losses are high and the training loss is initially close to 1.0 and then reduces. Validation loss also starts a bit larger than the training loss but also drops gradually. The two losses decrease at a very

high rate in the initial 20 epochs, and thereafter the rate decreases slowly. At epoch 100 the training loss has decreased to about 0 and the validation loss has leveled out at about 0.2 meaning that the model has been trained well on the training data and is generalizing satisfactorily on the validation data.

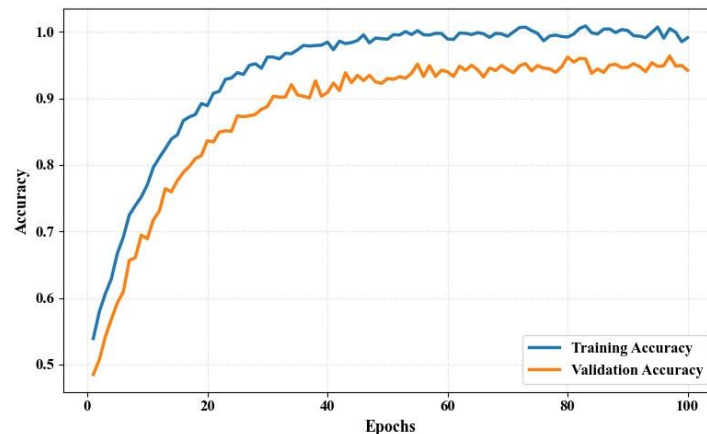


Figure 8. Training and validation accuracy

The accuracy of the training and validation during 100 epochs is depicted in Figure 8. At the beginning of the process, the training and validation accuracy are approximately 0.5. The training accuracy is also increasing gradually and it approaches almost 1.0 at epoch 100. The validation accuracy increases as well, yet it levels off to about 0.9 in the end of the training. The two curves have a strong increment in the initial 20 epochs and thereafter a steady increment. The accuracy of training is steadily increasing, and the accuracy of validation is no longer increasing, indicating that the model is doing well on training data and extrapolating relatively well to the validation data.

5. Conclusion and Future Scope

In conclusion, the PA-LSTM model that has been suggested is a significant step forward in classifying learning outcomes and self-efficacy levels of students. It is a model that understands the dynamics of student-AI interactions by capturing both temporal and semantic dependencies. The model's adaptability and personalized feedback capabilities make it very effective for real-time applications. The PA-LSTM model is more context-sensitive than traditional ones, therefore, it is a considerable tool for personalized learning in classroom situations. The present study lays down a strong foundation for the enhancement of student engagement and self-efficacy which, in turn, will lead to the creation of more powerful and responsive AI-based educational tools that are

still in their infancy. The outcomes of the experiment point to the high level of performance of the PA-LSTM model and significant success in the accuracy, precision, and recall. The model proposed had an accuracy of 0.975 to 0.995, precision of 0.97 to 0.99 and a recall of 0.97 to 0.992 with 10 k-folds. These findings prove the efficiency of the model in categorizing the learning outcomes and the self-efficacy level of students with almost perfect results. The research may also be extended to future studies with more varied interactions with the students and additional modalities, including facial expressions or physiological data, which would be used to provide the students with even more personalized feedback. Further, the investigation of the PA-LSTM model integration with other AI tools to assess and improve student learning in real-time will provide new prospects of the student learning experience.

Data Availability Statement

Data will be made available upon reasonable request.

Funding Statement

No funding available.

Conflict of Interest Disclosure

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Ethics Approval Statement

Not Applicable.

Permission to Reproduce Material from Other Sources

Not Applicable.

REFERENCES

- D. R. Serrano *et al.*, "Artificial intelligence (AI) applications in drug discovery and drug delivery: Revolutionizing personalized medicine," *Pharmaceutics*, vol. 16, no. 10, p. 1328, 2024.
- L. Cai, M. M. Msafiri, and D. Kangwa, "Exploring the impact of integrating AI tools in higher education using the Zone of Proximal Development," *Educ Inf Technol*, vol. 30, no. 6, pp. 7191–7264, Apr. 2025,
- W. Cain, "Prompting change: Exploring prompt engineering in large language model AI and its potential to transform education," *TechTrends*, vol. 68, no. 1, pp. 47–57, 2024.
- S. Avsec and D. Rupnik, "From Transformative Agency to AI Literacy: Profiling Slovenian Technical High School Students Through the Five Big Ideas Lens," *Systems*, vol. 13, no. 7, p. 562, Jul. 2025,
- C. Li and C. Zhang, "Exploring the Current Landscape of Primary School Physical Education Within the Framework of the New Curriculum Reform: A Quality Evaluation Model Perspective," *J Knowl Econ*, vol. 15, no. 4, pp. 20677–20698, Dec. 2024,
- D. Roy, "Pedagogical Restructuring of Business Communication Courses: AI-Enhanced Prompt Engineering in an EFL Teaching Context," in *Artificial Intelligence in Education: The Intersection of Technology and Pedagogy*, P. Ilic, I. Casebourne, and R. Wegerif, Eds., Cham: Springer Nature Switzerland, 2024, pp. 247–287.
- F. Sarshartehrani, E. Mohammadrezaei, M. Behravan, and D. Gracanin, "Enhancing E-Learning Experience Through Embodied AI Tutors in Immersive Virtual Environments: A Multifaceted Approach for Personalized Educational Adaptation," in *Adaptive Instructional Systems*, R. A. Sottilare and J. Schwarz, Eds., Cham: Springer Nature Switzerland, 2024, pp. 272–287.
- L. Zheng, C. Wang, X. Chen, Y. Song, Z. Meng, and R. Zhang, "Evolutionary machine learning builds smart education big data platform: Data-driven higher education," *Applied Soft Computing*, vol. 136, p. 110114, Mar. 2023,
- G. Maimaiti and K. F. Hew, "Gamified self-regulated learning improves EFL Reading comprehension, motivation, self-regulation skills and process patterns: Quasi-experiment with process mining," *The Internet and Higher Education*, vol. 67, p. 101042, Oct. 2025,
- V. G. A. Hakim, N. A. Paiman, and M. H. S. Rahman, "Genie-on-demand: A custom AI chatbot for enhancing learning performance, self-efficacy, and technology acceptance in occupational health and safety for engineering education," *Comp Applic In Engineering*, vol. 32, no. 6, p. e22800, Nov. 2024,
- P.-H. Li, T.-T. Wu, and Y.-M. Huang, "Improving inquiry-based learning through automated ChatGPT-based inquiry prompt and behavioral image recognition," *Educ Inf Technol*, Sep. 2025,
- B. Tang, J. Liang, W. Hu, and H. Luo, "Enhancing Programming Performance, Learning Interest, and Self-Efficacy: The Role of Large Language Models in Middle School Education," *Systems*, vol. 13, no. 7, p. 555, 2025.
- R. Azevedo, D. Dever, and M. Wiedbusch, "Enhancing Engineering Education: A Multimodal Approach to Personalization and Adaptation Using Artificial Intelligence in Game-Based Learning," *Advances in Artificial Intelligence Applications in Industrial and Systems Engineering*, pp. 281–306, 2025.
- Y.-J. Ju *et al.*, "Developing the questionnaire of self-efficacy and needs in using large-language model-based AI services," *Curr Psychol*, vol. 44, no. 9, pp. 8158–8176, May 2025,
- J. Kong, J. Liu, G. Chen, and W. Shang, "Assessing AI literacy in college students: the mediating role of self-efficacy in motivational commitment pathways," *Educ Inf Technol*, Aug. 2025,