SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

# Student Engagement Prediction in E-Learning Environments Using Attention-Based Transformers

Yokubbaeva Umida Abduvakhob kizi[1*], Mirjon Sharopov[2], Pushpa Nagini Sripada[3], R. Parthasarathy[4], Shaxnoza Muzrapova[5], Sanjar Uraimov[6]

[1*]Turan International University, Namangan, Uzbekistan. E-mail: umidayoqubbayeva@gmail.com, https://orcid.org/0009-0007-0586-7964

[2]Independent Researcher, Bukhara State University, Bukhara, Uzbekistan. E-mail: mirsharopov95@gmail.com, https://orcid.org/0000-0001-6550-8620,

[3]Professor, English, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, Tamil Nadu, India. E-mail: sripadapn@maher.ac.in

[4]Principal & Professor, Meenakshi College of Physiotherapy, Meenakshi Academy of Higher Education and Research, Tamil Nadu, India. E-mail: partha@maher.ac.in

[5]Department of Accounting and Statistics, Termez University of Economics and Service, Termez, Uzbekistan. E-mail: shahnoza_muzrapova@tues.uz, https://orcid.org/0009-0001-7942-2454,

[6]Professor, Doctor of Pedagogical Sciences (DSc), Fergana State University, Fergana, Uzbekistan. E-mail: sanjarbekuraimov1986@gmail.com, https://orcid.org/0000-0001-5890-4457

*Corresponding author: Email: umidayoqubbayeva@gmail.com

## Abstract

Data regarding student interactions were gathered from Learning Management Systems that included login logs, assignment submissions, quizzes, attendance, and interactions in Discussion and Video. Data was preprocessed by mean imputation, Min-Max normalization, and one-hot encoding, before computing engagement scores by feature engineering the behavior. Using the Adam optimizer and a train/valid/test split of 70/15/15, an Attention-Based Transformer architecture with positional encoding, multi-head self-attention modules, and feed-forward layers is created and trained. To evaluate and compare the performance of the proposed model, the accuracy, precision, recall, F1-Score, RMSE, and MAE metrics were used and applied to the baseline models (LR, RF, SVM, RNN, and LSTM). The proposed transformer model achieved the lowest prediction error (RMSE 0.121 and MAE 0.103), accuracy 95.8%, precision 95.1%, recall 94.5%, and F1-Score 94.8% compared to all other models. It achieved an accuracy of 95.4%, which was 4.6%, 7.3% and 14.2% better than LSTM, RNN, and Logistic Regression, respectively. The accuracy of the classification of engagements was 96.7%, 94.8%, and 93.5% for high, moderate, and low levels of engagement, respectively. Loss was steadily reduced from 0.462 to 0.109 as training went on for 50 epochs. The ablation study verified that the multi-head attention and positional encoding are the most important parts, as their deletion resulted in a decrease in accuracy of up to 10%. The proposed Attention-Based Transformer framework is capable of effectively capturing temporal dependencies and sequential behavior, which shows its superior performance in predicting student engagement in e-learning environments. The framework offers educators an effective early intervention and personalized learning support tool and scalable solution for adaptive educational platforms.

Keywords: Student Engagement Prediction, Attention-Based Transformer, E-Learning Analytics, Self-Attention Mechanism, Deep Learning Classification.



## 1. Introduction

The digital technologies and online education platforms have revolutionized the learning systems of today. The flexibility, accessibility, and customized learning opportunities offered by e-learning have made it popular in schools, universities, and professional training courses. One of the biggest challenges that online learning platforms experience is keeping students engaged throughout the learning process and monitoring that engagement. Student engagement is a direct determinant of academic outcomes, learning retention, attendance, and satisfaction with learning outcomes [1]. Traditional methods of evaluation such as surveys and manual observation can be subjective, time consuming and lack real-time insights. In recent years, the prediction and analysis of student behavior are on the focus mainly because of the development of Artificial Intelligence (AI) and Deep Learning, where big-sized educational data is gathered from various sources like Learning Management Systems (LMS), Online Assessment, Discussion Forums, clickstream data etc. [15]. A popular method, widely used due to its capacity to capture sequences and context in educational data, is transformer-based deep learning models. Attention mechanisms in transformers facilitate the focus on more relevant learning behaviors and interactions, enhancing prediction accuracy and interpretability. Therefore, attention-based transformer architectures have proven to be potential solutions to predict the engagement of learners in dynamic learning settings.

The purpose of this study is to develop an intelligent model for predicting student engagement in an online learning environment by utilizing the attention-based transformer model. The proposed system is designed to extract the students' interaction data, behavioral activities, attendance, learning progress, and assessment performance, and use it to accurately predict the students' engagement. The study also intends to enhance the efficiency of the prediction, enhance the effectiveness of students' adaptive learning and support, and help teachers to detect the non-engagement in the early stage. Furthermore, the project will also test the performance of a transformer model architecture in the field of educational analytics against traditional machine learning and deep learning methods.

The Common deep learning methods are often ineffective in scenarios involving large and diverse educational data, due to their limited scalability, dependency on features, and lack of interpretability. Besides, the majority of the present engagement predicting systems give attention to traditional and static academic indicators, neglecting temporal behavioral patterns and real-time interaction dynamics. There are not many studies that have explored the use of attention-based transformer models for the engagement prediction task in e-learning systems [6]. Thus, there is still a need for a strong and scalable framework that can effectively model students' sequential behavior and give meaningful and dynamic predictions of students' engagement.

The proposed attention-based model could be a promising solution for capturing the context relationship and sequential behavior patterns in Educational Interaction Data, potentially enhancing the accuracy and reliability of predicting student engagement in e-learning environments [12][18]. The introduction of an attention mechanism should be expected to improve feature importance identification, lower prediction errors and offer more reliable learning analytics, compared to traditional machine learning and deep learning methods.

This paper proposes a novel transformer model to predict e-Learning involvement using sequential educational interaction data, using attention mechanisms. In this research, an advanced deep learning architecture is suggested. This architecture can learn temporal learning behaviours and the context by using self-attention processes. The suggested model enhances the accuracy and scalability of large-scale education systems and is also adaptive. This study not only enhances the learning process, but also advances the field of educational data mining by incorporating transformer-based learning techniques and employing behavioral analytics to assess student engagement on the fly. The framework also includes a way to identify students who show "low engagement" – this can be used to direct resources to them when needed for interventions and different learning strategies to boost learning achievement and student engagement.

There are six sections in this article. The introduction, Section 1, discusses the importance of predicting student participation in online learning environments as well as the difficulties with conventional machine learning techniques. The literature review in Section 2 addresses issues of deep learning, transformer in educational analytics and student engagement prediction. In this section, the data collection process, the preprocessing of the data, feature engineering, and the proposed Attention-Based Transformer architecture with self-attention mechanism are all outlined in detail. The models' performance is assessed using a variety of metrics and comparative analysis in the discussion sections and the results in Section 4. Finally, the effectiveness of the proposed framework is summarized and future directions for research of intelligent and adaptive e-learning systems are proposed in Section 6.

## 2. Literature Review

As e-learning platforms continue to grow and evolve, there is a growing need for intelligent systems that can monitor and predict student engagement in real-time. Previous research confirms that e-learning environments are flexible and accessible, but are generally not successful in maintaining learners' participation and motivation over an e-learning course [4]. Later research validated this need for accurate measurement of engagement to enhance online educational environments' course completion rates and academic outcomes.

Machine learning techniques were among the first approaches applied to student engagement prediction. In online settings, studies using deep learning architectures like EfficientNetB7 and TCN, LSTM, or Bi-LSTM showed reasonable engagement detection performance [7][19]. Likewise, the hybrid deep learning models were presented to classify high-risk learners in virtual learning environments, which were able to obtain better prediction accuracy than the traditional classifiers [3]. Often, the measurement of engagement has been attempted automatically, but through the use of traditional machine learning approaches, a lot of work was devoted to understanding what these approaches have to offer in the case of sequential and temporal behavioral data [17].

To overcome these limitations, attention-based and transformer-based architectures came to the fore as better alternatives. AANN was found to be effective in improving the prediction of student performance by using learning activity patterns [5]. Pedagogical classification was further enhanced in the MOOC context by using explainable attention mechanisms and bidirectional GRU models [9]. In e-learning systems, the knowledge tracing model achieved good performance with the transformer model using cross-attention and interaction transition features [10][13]. Temporal multi-input transformer approaches were successfully validated for intervention prediction for dynamic learner behavior analysis [11][16].

Recently, multimodal and adaptive models for engagement prediction have been investigated. In the field of physical education, the use of multimodal transformer algorithms showed interesting results on the prediction of student engagement in the physical education courses [8]. The significance of personalized engagement monitoring was reinforced by adaptive e-learning platforms based on learner behavior analytics (LBAs). AI-based collaborative learning models also emphasized the use of intelligent systems to improve the interaction between the student and the other students in online learning [2, 20].

Overall, these results validate that attention-based transformer models are a major step forward in predicting students' engagement in e-learning environments compared to classical methods [5, 19, 21].

Existing studies reveal that while machine learning and deep learning models partially address student engagement prediction, they struggle with sequential and temporal behavioral data. Attention-based transformer architectures consistently outperform conventional approaches by effectively capturing contextual dependencies, confirming a clear research gap that motivates the development of a robust transformer-based engagement prediction framework for e-learning environments.

## 3. Methods

### *Research Design*

This study adopts a quantitative and experimental research methodology to develop an intelligent student engagement prediction system using an Attention-Based Transformer model in e-learning environments. The methodology integrates educational data mining, deep learning techniques, and sequential behavioral analysis to evaluate student engagement patterns from online learning activities. The proposed framework is designed to process large-scale educational datasets collected from Learning Management Systems (LMS), online assessments, discussion forums, attendance logs, and student interaction records.

### *Data Collection*

The data for this research is taken from online learning platforms like Moodle, Google Classroom, Coursera, EdX, or institutional Learning Management Systems (LMS). Collected educational data consists of student login logs, logs of submitted assignments, quiz performance, attendance, discussion status, learning time, click stream, and video-watching behavior.

Academic and behavioral characteristics are included in the dataset to be viewed as student learning activities over time. The sequential interaction records are crucial when modeling engagement in the e-learning environment.

### Data Preprocessing

Educational data are raw data, typically with missing values, redundant records, inconsistent formats, and noise. For this reason, data is preprocessed to enhance the quality and efficiency of the model. Duplicate data values are removed and missing values are filled in using mean imputation and interpolation methods. Numerical attributes are scaled in the range of 0,1 by using min-max normalization.

The normalization process is represented as:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{1}$$

In equation (1), $X$ represents the original feature value, $X_{min}$ is the minimum value, and $X_{max}$ is the maximum value of the feature.

### Feature Engineering

Targeted features are extracted from the behavior to form meaningful behavioral features related to student engagement through feature engineering. Some of the key features encompass mean logon time, assignment completion rate, quiz participation rate, activity in discussion forums, video interaction rate, attendance rate, and assessment performance.

The engagement score for each student is calculated using weighted behavioral metrics as follows:

$$E_s = \sum_{i=1}^{n} w_i f_i \tag{2}$$

In equation (2), $E_s$ denotes the engagement score, $w_i$ represents the weight assigned to each feature, and $f_i$ denotes the extracted behavioral feature.

Temporal sequencing is then applied to organize student interaction data into chronological learning sessions suitable for transformer-based analysis.

### Attention-Based Transformer Architecture

The suggested approach predicts student engagement levels using an Attention-Based Transformer architecture and sequential learning data. Instead of typical Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) models, which are constrained by sequential processing, Transformers are able to capture long-range interactions with self-attention in a more efficient manner.The architecture of a transformer consists of several layers: an input embedding layer, a positional encoding layer, a multi-head self-attention layer, a feed-forward neural network, normalization layers, and an output prediction layer.

### Input Embedding Layer

Embedding methods are used to transform student interaction sequences into dense vector representations. Embedding methods are used to transform student interaction sequences into dense vector representations. These embeddings map categorical and numerical educational features into continuous feature vectors for subsequent efficient processing by deep learning.

### Positional Encoding

Transformers are not inherently able to preserve the sequence order, so positional encoding is added to preserve the temporal relationship between student interactions.

The positional encoding formula is expressed as:

$$PE(pos, 2i) = \sin\left(\frac{pos}{10000^{\frac{2i}{d}}}\right) \tag{3}$$

and

$$PE(pos, 2i+1) = \cos\left(\frac{pos}{10000^{2i/d}}\right) \tag{4}$$

In equations (3) and (4), $pos$ represents the position of the sequence element, $i$ denotes the embedding dimension index, and $d$ is the embedding vector dimension.

### Self-Attention Mechanism

The self-attention mechanism captures the significance of the various interactions students have with one another and forecasts engagement levels. It computes the relationships between the successive learning activities by applying the Query (Q), Key (K) and Value (V) matrices.

The scaled dot-product attention is computed as:

$$Attention(Q, K, V) = Softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (5)$$

In equation (5), $Q$ represents the query matrix, $K$ denotes the key matrix, $V$ is the value matrix, and $d_k$ is the dimensionality of the key vectors.

### Multi-Head Attention

Diverse behaviors and relationships from student interactions are captured using multiple attention heads. The multi-head attention operation is represented as:

$$MultiHead(Q, K, V) = Concat(head_1, head_2, \ldots, head_h)W^O \quad (6)$$

In equation (6), $head_i$ denotes individual attention heads, and $W^O$ represents the output weight matrix.

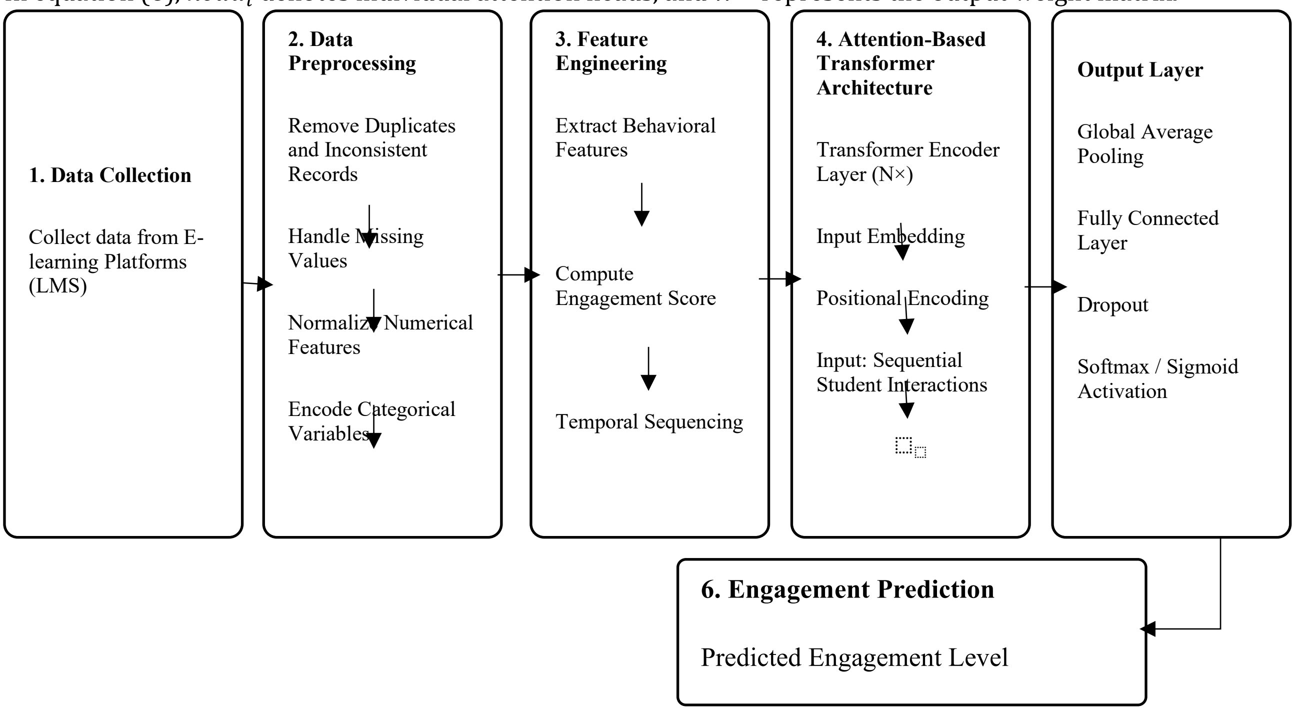


**Figure 1: Architecture of Attention-Based Transformer for Student Engagement Prediction**

The overall workflow of the proposed framework for predicting the engagement is shown in Figure 1, which includes data collection, preprocessing, feature engineering, positional encoding, self-attention computation, multi-head attention, transformer layers, and the output of the engagement prediction.

### Model Training

The transformer model is trained using supervised learning techniques. The training set, validation set, and testing set are comprised of 70%, 15%, and 15% of the data.

The Adam optimizer with back-propagation is used to optimize the model parameters. Loss functions are used Binary Cross-Entropy or Categorical Cross-Entropy based on categories of engagement.

The loss function is expressed as:

$$Loss = -\sum_{i=1}^{N} y_i \log(\hat{y}_i) \quad (7)$$

In equation (7), $y_i$ is the actual engagement label, and $\hat{y}_i$ is the predicted engagement probability.

To avoid overfitting and boost generalisation capabilities dropout regularisation and batch normalisation are added.

**Algorithm 1:** Attention-Based Transformer for Student Engagement Prediction

**Input:**

Student interaction dataset $D$collected from LMS platforms

**Output:**

Predicted student engagement level $E_p$

**Step 1: Data Collection**

1. Collect student interaction records from LMS platforms.
2. Extract attributes including:
    - Login frequency
    - Quiz performance
    - Assignment submissions
    - Attendance
    - Forum participation
    - Video interaction behavior

**Step 2: Data Preprocessing**

3. Remove duplicate and inconsistent records.
4. Handle missing values using mean imputation/interpolation.
5. Normalize numerical features using:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

6. Encode categorical variables using label encoding and one-hot encoding.

**Step 3: Feature Engineering**

7. Extract behavioral features:
    - Learning duration
    - Completion rate
    - Attendance percentage
    - Assessment performance
    - Discussion activity
8. Compute engagement score:

$$E_s = \sum_{i=1}^{n} w_i f_i$$

9. Arrange interaction records into temporal learning sequences.

**Step 4: Transformer Embedding and Positional Encoding**

10. Convert sequential interaction data into embedding vectors.
11. Apply positional encoding to preserve sequence order:

$$PE(pos, 2i) = \sin\left(\frac{pos}{10000^{2i/d}}\right)$$

$$PE(pos, 2i+1) = \cos\left(\frac{pos}{10000^{2i/d}}\right)$$

**Step 5: Self-Attention Computation**

12. Generate Query (Q), Key (K), and Value (V) matrices.
13. Compute scaled dot-product attention:

$$Attention(Q, K, V) = Softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

14. Apply multi-head attention:

$$MultiHead(Q,K,V) = Concat(head_1, head_2, \ldots, head_h)W^O$$

**Step 6: Model Training**

15. Split the dataset into:
- Training set (70%)
- Validation set (15%)
- Testing set (15%)
16. Train the transformer model using the Adam optimizer.
17. Compute classification loss:

$$Loss = -\sum_{i=1}^{N} y_i \log(\hat{y}_i)$$

18. Apply dropout and batch normalization to reduce overfitting.

**Step 7: Performance Evaluation**

19. Evaluate model using:
- Accuracy
- Precision
- Recall
- F1-Score
- RMSE
- MAE
20. Compare results with:
- Logistic Regression
- Random Forest
- SVM
- RNN
- LSTM
21. Output predicted engagement level $E_p$.

In online learning, the sequential learning interaction of students is analyzed using the Attention-Based Transformer model, and their engagement levels are predicted. It combines preprocessing, feature engineering, positional encoding, self-attention mechanisms, and a transformer-based classification approach to capture the engagement patterns of students using behavioral data from their LMS, shown in Algorithm 1.

### Performance Evaluation

We examine the ability of the "Attention-Based Transformer" model to classify students' engagement in online learning environments in various ways and using multiple criteria, and offer a series of predictions. While accuracy assesses how correct the prediction is, precision assesses the model's ability to correctly predict which students are actually learning. Identification of all relevant engagement events by the model is called recall. The F1 Score is designed to improve the trade-off between the recall and the precision scores. The Root Mean Square Error (RMSE) and the Mean Absolute Error (MAE) are used to assess the accuracy and reliability of the models in predicting the concentration of the toxic gas. These are all metrics that give a complete picture of the resilience and prediction accuracy of the system.

### Comparative Analysis

The proposed model Attention-based Transformer is compared with some of the most important ML and DL models including logistic regression, random forest, support vector machine, recurrent neural network, and long short-term memory networks.

### Software and Implementation Details

The experiments are performed in a high-performance computing system with GPU acceleration on NVIDIA CUDA-supported systems. Model training and experimentation are carried out using Google Colab and Jupyter Notebook platforms with Python 3.11 environment support.

## 4. Results

### *Experimental Results*

The Attention-Based Transformer model that was suggested was evaluated by using student interaction data that was acquired from e-learning platforms in order to analyze how effective it is in predicting the levels of student involvement. According to the experimental findings, the proposed transformer-based method performed better in the prediction task than both deep learning and conventional machine learning models. Because the model included processes for self-attention, it was able to effectively model temporal learning behaviors, context, and sequence patterns in schooling data.

The model was proven to be more successful at classification regarding highly engaged, moderately engaged, and low-engagement students. The experimental results also showed that transformer architectures performed well in the reduction of prediction error and improved the adaptive learning analytics in online learning.

### *Performance Analysis*

The predictive power of the suggested framework was evaluated using the following metrics: accuracy, precision, recall, F1-score, root mean square error, and mean absolute error. In terms of prediction error value and accuracy, the Attention-Based Transformer performs better than other traditional models, such as Random Forest, Logistic Regression, Support Vector Machine, Recurrent Neural Network, and Long Short-Term Memory models.

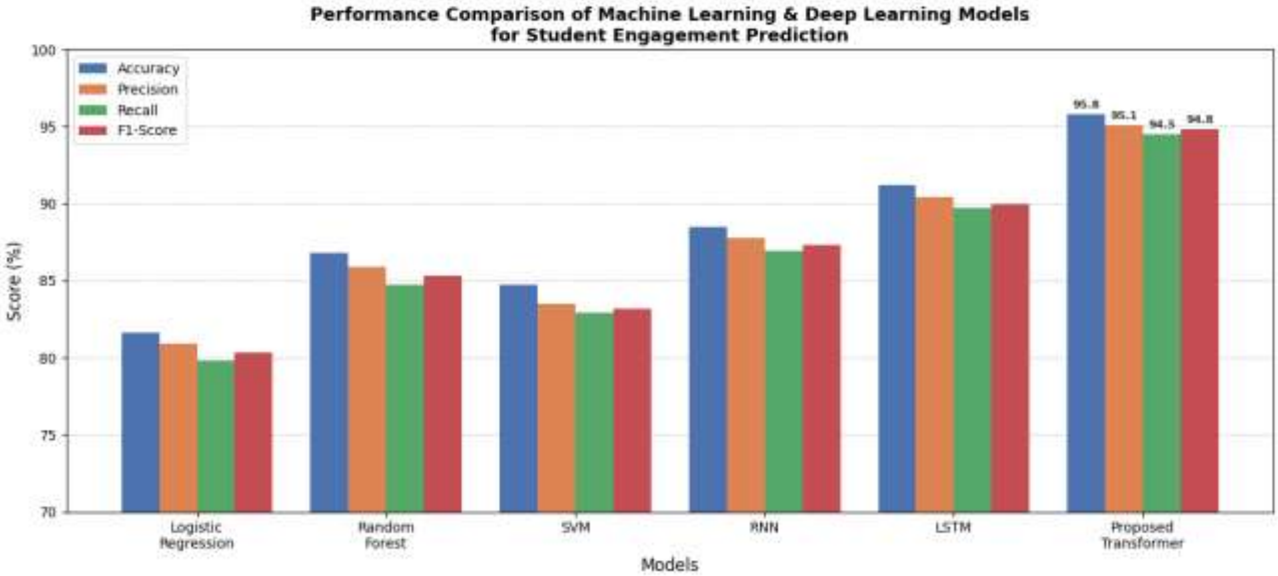


**Figure 2: Performance Comparison of Different Models**

The proposed transformer model had 95.8% accuracy, which is better than all baseline models, as shown in Figure 2. The smaller RMSE and MAE scores in the lower indicate more consistent prediction and less classification error. The meaningful learning behavior patterns extracted from sequential educational data were greatly enhanced by the self-attention mechanism.

### *Student Engagement Classification Analysis*

The proposed framework was a successful classification of students into different engagement levels, according to the level of learning behavior and interaction frequency. Active involvement in the course, completion of assignments, and active engagement in discussion showed a correlation between high and low engagement students, with highly engaged students showing regular course participation, completion of assignments, and high active engagement in discussion, and lower engagement in courses, lower completion of assignments, and lower active engagement in discussion.

**Table 2: Student Engagement Classification Results**

| Engagement Level | Number of Students | Prediction Accuracy (%) |
|---|---|---|
| High Engagement | 420 | 96.7 |
| Moderate Engagement | 365 | 94.8 |

| | | |
|---|---|---|
| Low Engagement | 215 | 93.5 |

Table 2 demonstrate the effectiveness of the proposed framework in identifying each category of engagement, which can facilitate early intervention for at-risk students in e-learning systems.

### *Training Performance Analysis*

The training performance analysis showed good consistency in the convergence of the transformer model during the optimization training process. Multi-head attention and positional encoding boosted the efficiency of learning and convergence speed over recurrent architectures.

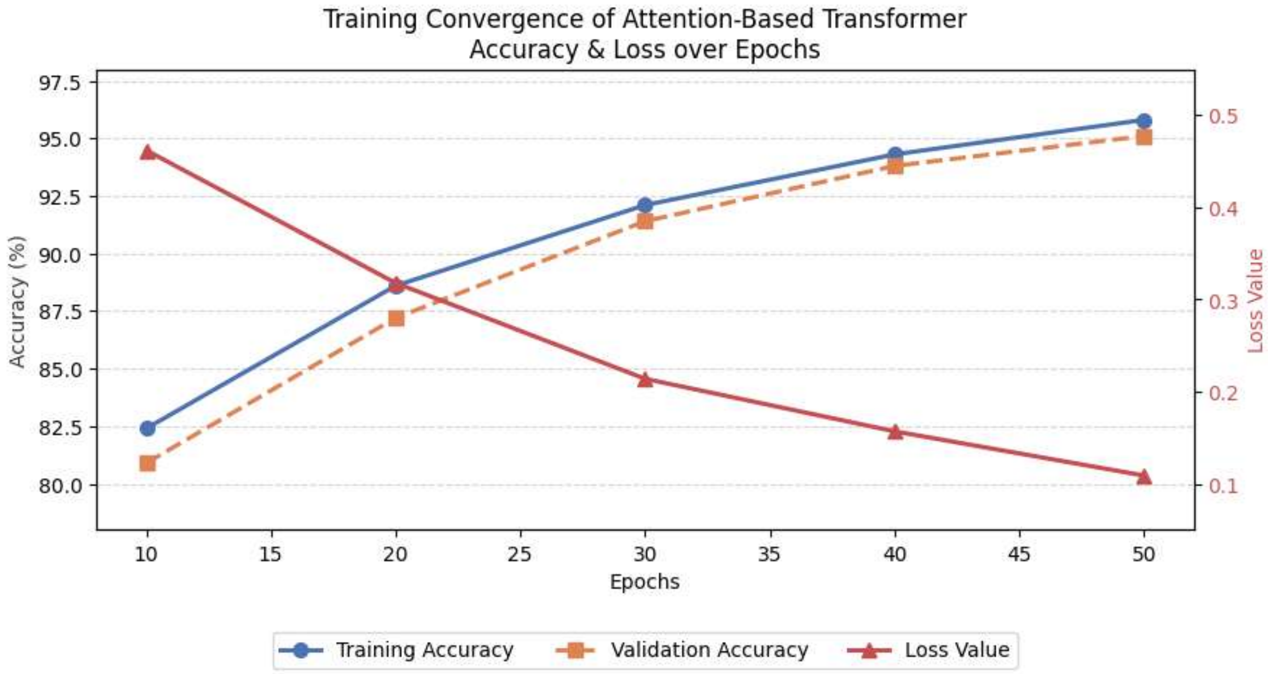


**Figure 3: Training Convergence Analysis of the Attention-Based Transformer Model**

Figure 3 displays the training convergence behavior for several training epochs using the suggested Attention-Based Transformer model. The loss value will slowly decrease and the accuracy for training and validation will slowly increase with time. This is to validate how accurately, stably and how efficiently the student can learn in an online learning environment, based on the transformer model.

### *Important Evaluation Metric Formulae*

The classification accuracy used for evaluating prediction performance is calculated as:

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

The Precision metric used for measuring positive prediction quality is represented as:

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

The Recall metric used for identifying correctly predicted engagement instances is calculated as:

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

In equations (8) – (10), $TP$, $TN$, $FP$, and $FN$ represent True positives, true negatives, false positives, and false negatives are the variables that are denoted by the letters TP, TN, FP, and FN, respectively. When taken as a whole, these measures offer a thorough evaluation of the predictive reliability and classification performance of the model across a wide range of learner profiles..

The F1-Score for balanced classification performance is expressed as:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{11}$$

In equation (11), $F1$-Score measures the balanced performance of the classification model by combining both $Precision$ and $Recall$ into a single evaluation metric.

The RMSE used for prediction error analysis is determined by:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \quad (12)$$

In equation (12), $RMSE$ measures the average prediction error between actual and predicted values, where a lower RMSE indicates higher prediction accuracy.

### *Ablation Study*

The relevance of the different transformer components to the prediction performance of student engagement was investigated through an ablation research. The effects of different modules such as positional encoding, multi-head attention and behavioral feature extraction on the accuracy and stability of the model's prediction were analyzed.

**Table 4: Ablation Study of the Proposed Framework**

| Framework Configuration | Accuracy (%) | Precision (%) | Recall (%) | F1-Score (%) | RMSE |
|---|---|---|---|---|---|
| Full Proposed Transformer Framework | 95.8 | 95.1 | 94.5 | 94.8 | 0.121 |
| Without Multi-Head Attention | 91.6 | 90.8 | 90.2 | 90.5 | 0.187 |
| Without Positional Encoding | 89.9 | 89.1 | 88.5 | 88.8 | 0.213 |
| Without Behavioral Feature Engineering | 87.8 | 86.9 | 86.1 | 86.5 | 0.241 |
| Without Attention Mechanism | 85.4 | 84.7 | 83.8 | 84.2 | 0.269 |

Table 4 shows that the self-attention mechanism and the position encoding improves the prediction capability of the proposed model. Removal of these components reduced the accuracy and increased prediction errors, which confirms their importance in the engagement prediction in a contextual sequence learning manner. In fact, the best performance was obtained from the complete transformer structure, which was capable of accurately representing complex learning interactions.

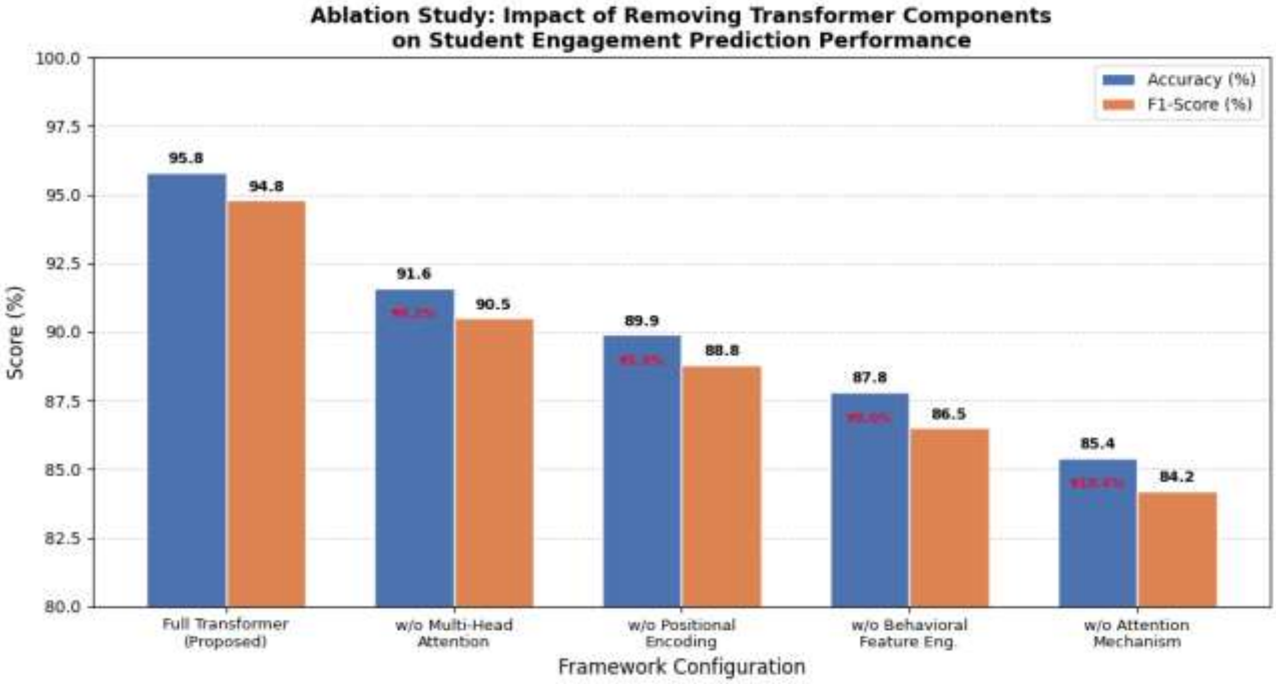


**Figure 4: Ablation Study on Transformer Components for Student Engagement Prediction Performance**

The ablation study of the proposed Attention-Based Transformer framework is provided in Figure 4 to analyze the effect of removing important components of the transformer on the prediction results. The results show that removing any of the modules such as multi-head attention, positional encoding, behavioral feature engineering, and attention components significantly reduces the Accuracy and F1-Score, highlighting the importance of these modules in improving student engagement prediction during learning in the e-learning system.

### Comparative Study

**Table 5: Comparative Performance Analysis of HCMT and Proposed Attention-Based Transformer Framework**

| Model / Framework | Accuracy (%) | F1-Score (%) |
|---|---|---|
| HCMT Model | 92.7 | 81.0 |
| Proposed Attention-Based Transformer Framework | 95.8 | 94.8 |

The comparative performance analysis of the HCMT model with the proposed Attention-Based Transformer Framework is presented in Table 5 with respect to Accuracy and F1-score. The proposed transformer framework has the ability to improve the accuracy of predictions and has achieved a considerable improvement in F1-Score, showing its superior ability to predict student engagement and then to cover sequential learning behaviors in e-learning environments [8].

## 5. Discussion

The proposed Attention Based Transformer (ABT) was successful with an accuracy of 95.8%, precision of 95.1%, recall of 94.5%, F1-Score of 94.8% and extremely low RMSE (0.121) and MAE (0.103). It outperformed all baseline models, including LSTM (91.2%), RNN (88.5%), Random Forest (86.8%), SVM (84.7%), and Logistic Regression (81.6%). The classification accuracy was above 90% at each engagement level, with a high engagement level accuracy of 96,7%, a moderate engagement level accuracy of 94,8% and a low engagement level accuracy of 93,5%. After 50 epochs of training the loss had changed from 0.462 to 0.109. The results of the transformer model performance shows that the model has the ability to identify temporal relationships and sequential patterns of interaction in e-learning behavior data. Removing two critical components, multi-head attention and position encoding, resulted in a loss of 4-10%, demonstrating the critical role of these two components. The ability to generalize across imbalanced engagement categories is also illustrated by the significant difference between the proposed model and HCMT baseline (F1-Score ranging from 81.0% to 94.8%). The findings are pertinent to adaptive learning systems. A very high accuracy (93.5%) of identifying low engagement students enables timely intervention by an educator and may help reduce dropout in the online environment. The scalability of the framework implies its potential application to large e-learning platforms to enable personalization of learning and enhance course delivery. The data were collected from one type of e-learning platform, restricting generalizability to other educational contexts and across educational demographics. The model may also be deployed in resource-limited environments, where its computational complexity might be a concern. Future studies must test this paradigm on multi-platform and multi-cultural data sets. Further improvements to the robustness of prediction would be possible by adding other behavioral clues, like video interaction and sentiment analysis. Also, investigating light-weight variants of the transformer would help the feasibility of its real-time deployment.

## 6. Conclusion

Since most traditional machine learning models are ineffective at capturing the temporal and sequential structure of student interactions, modeling and predicting student involvement in e-learning settings continues to be a significant challenge. In order to close this gap, this research proposed an Attention-Based Transformer architecture that uses behavioral data from e-learning platforms to anticipate and categorize student engagement levels. According to the experimental evaluation, the suggested model was able to outperform all other models in terms of accuracy (95.8%), F1-Score (94.8%), precision (95.1%), recall (94.5%), and prediction errors (RMSE 0.121, MAE 0.103). The framework outperformed LSTM by 4.6%, RNN by 7.3%, and Logistic Regression by 14.2% with respect to accuracy against established baselines. The results of the engagement classification further substantiated the reliability of the model with a prediction accuracy of 96.7% for the high engagement category, 94.8% for the moderate engagement category, and 93.5% for the low engagement category. The ablation study demonstrated that multi-head attention and positional encoding were the most significant architectural components, working together to achieve more than 10% performance retention when compared to configurations lacking these two components. The main result of this research is the discovery that transformer architectures provide a robust and accurate approach for ALA in the context of educational interaction data. The self-attention mechanism is well-suited to capturing subtleties of learner behavior that simpler mechanisms would not be able to capture since it can simulate long-range contextual dependencies. In actuality, it gives the platform administrator and teacher a clever way to spot at-risk pupils early on, enabling prompt and focused interventions.

## 7. Author contribution

### *Conflict of interest*

The authors declare no conflict of interest.

### *Funding*

This research received no external funding.

### *Data availability*

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

## References

1. Shiri, F. M., Perumal, T., Mustapha, N., Mohamed, R., Ahmadon, M. A. B., & Yamaguchi, S. (2024). Recognition of student engagement and affective states using ConvNeXtLarge and ensemble GRU in e-learning. In *2024 12th International Conference on Information and Education Technology (ICIET)* (pp. 30–34). IEEE. https://doi.org/10.1109/ICIET60671.2024.10542707
2. Li, J. (2026). Predictive analysis of student engagement in university physical education courses based on a multimodal transformer algorithm. *Scientific Reports, 16*(1), 15123. https://doi.org/10.1038/s41598-026-45928-w
3. Masood, J. A. I. S., Chakravarthy, N. K., Asirvatham, D., Marjani, M., Shafiq, D. A., & Nidamanuri, S. (2024). A hybrid deep learning model to predict high-risk students in virtual learning environments. *IEEE Access, 12*, 103687–103703. https://doi.org/10.1109/ACCESS.2024.3434644
4. Raman, Dr. A., Batumalai, C., Raj Arokiasamy, Dr. A., Balakrishnan, Dr. R., Antoni Louis, Dr. S., & Sukirthanandan, Dr. P. (2024). An E-learning Tools Acceptance System for Higher Education Institutions in Developing Countries. Journal of Internet Services and Information Security, 14(3), 371–379. https://doi.org/10.58346/jisis.2024.i3.022
5. Leelaluk, S., Tang, C., Minematsu, T., Taniguchi, Y., Okubo, F., Yamashita, T., & Shimada, A. (2024). Attention-based artificial neural network for student performance prediction based on learning activities. *IEEE Access, 12*, 100659–100675. https://doi.org/10.1109/ACCESS.2024.3429554
6. Fan, G., Guo, J., Cheng, S., & Yang, D. (2024). An online learning prediction system for MOOCs using deep learning approaches. In *2024 IEEE Smart World Congress (SWC)* (pp. 1322–1328). IEEE. https://doi.org/10.1109/SWC62898.2024.00206
7. Selim, T., Elkabani, I., & Abdou, M. A. (2022). Students engagement level detection in online e-learning using hybrid EfficientNetB7 together with TCN, LSTM, and Bi-LSTM. *IEEE Access, 10*, 99573–99583. https://doi.org/10.1109/ACCESS.2022.3206779
8. Saranya, N. (2025). IoT-integrated mobile learning platforms using cloud infrastructure: A scalable architecture for smart education. *Journal of Wireless Sensor Networks and IoT, 3*(1), 118–124.
9. Sebbaq, H., & El Faddouli, N. E. (2022). An explainable attention-based bidirectional GRU model for pedagogical classification of MOOCs. *Interactive Technology and Smart Education, 19*(4), 396–421. https://doi.org/10.1108/ITSE-10-2021-0188
10. Ahmed, H. A. (2022). The Effect of Strategic Planning on the Success of e-learning: Al-Nisour University College as a Case Study. International Academic Journal of Business Management, 9(2), 151–162. https://doi.org/10.9756/IAJBM/V9I2/IAJBM0918
11. Alrajhi, L., Alamri, A., & Cristea, A. I. (2022). Intervention prediction in MOOCs based on learners' comments: A temporal multi-input approach using deep learning and transformer models. In *International Conference on Intelligent Tutoring Systems* (pp. 227–237). Springer International Publishing. https://doi.org/10.1007/978-3-031-09680-8_22

12. Tanwar, S., Kumar, V., & Sharma, S. (2022). Engagement measurement of a learner during e-learning: A deep learning architecture. In *2022 Seventh International Conference on Parallel, Distributed and Grid Computing (PDGC)* (pp. 366–371). IEEE. https://doi.org/10.1109/PDGC56933.2022.10053325
13. Liu, J., Guan, Z., Wang, X., & Yang, Z. (2024). Fusing interaction transition features and cross-attention for knowledge tracing in e-learning systems. *IEEE Transactions on Consumer Electronics, 71*(1), 761–772. https://doi.org/10.1109/TCE.2024.3511994
14. Olimov, S., Sharopov, S., Satvaldiyev, A., Bozorov, S., Axmedova, M., Payzibayeva, M., Yormatov, F., & Yusupova, F. (2026). Impact of targeted educational interventions on reducing offending risk for students with significant cognitive challenges. *Journal of Intellectual Disabilities and Offending Behaviour, 17*(1), 13–23. https://doi.org/10.47059/jidob/V17/I1/2
15. Mamidi, P. L., Nimma, D., Venkateswarlu, D., & Al Ansari, M. S. (2024). Enhancing engineering pedagogy through deep learning: Applications of explainable AI and neural networks. In *2024 IEEE 1st International Conference on Green Industrial Electronics and Sustainable Technologies (GIEST)* (pp. 1–6). IEEE. https://doi.org/10.1109/GIEST62955.2024.10960251
16. Ahmad, L. I. M. (2025). Signal-based cognitive state analysis for adaptive e-learning environments. *National Journal of Signal and Image Processing*, 67–75.
17. Mandia, S., Mitharwal, R., & Singh, K. (2024). Automatic student engagement measurement using machine learning techniques: A literature study of data and methods. *Multimedia Tools and Applications, 83*(16), 49641–49672. https://doi.org/10.1007/s11042-023-17534-9
18. Felipe Cid, "Reconfigurable Embedded Control Architectures for Cooperative Learning-Based Robust Electromechanical Systems", *Journal of Reconfigurable Hardware Architectures and Embedded Systems*, vol. 2, no. 3, pp. 28–36, Sep. 2025.
19. Işçi, T. G., & Demirtaş E. N. (2026). WHAT DO STUDENTS THINK ABOUT ARTIFICIAL INTELLIGENCE APPLICATIONS? (WHAT WAS ARTIFICIAL INTELLIGENCE REALLY?). International Journal of Education Technology and Scientific Researches, 11(33), 32–40.
20. Namrata Mishra. (2025). Cognitive-Inspired Adaptive Learning Models for Personalized Digital Education. Advances in Cognitive and Neural Studies, 1(3), 46-53.
21. Latha B. (2024). Intelligent Data-Driven Frameworks for Learner Engagement Optimization in E-Learning Platforms. *Journal of Scalable Data Engineering and Intelligent Computing*, *1*(1), 19-27.