Iraqi Journal for Applied Science (IJAS)
Vol. 3, No. 1, March 2026, pp. 74-83
ISSN: 3006-5828, DOI: 10.69923/mewnft50

Adaptive Learning with Attention-Based Knowledge Tracing and Risk Prediction for Improved Student Outcomes

Hazim Noman Abed
Department of Computer Science, College of Science, University of Diyala, Diyala, Iraq.

Article Info

Article history:

Received: 11, 01, 2026
Revised: 30, 02, 2026
Accepted: 20, 03, 2026
Published: 30, 03, 2026

Keywords:

Adaptive learning,
Knowledge tracing,
Attention mechanism,
Contextual bandit,
educational data mining,
Fairness,
Personalized tutoring.

ABSTRACT

The objective of adaptive learning environments is to offer personalized learning experiences, and many current systems are constrained by fixed instructional strategies, noisy learner-interaction data, and a lack of attention to equity across diverse student groups. In this paper, a proposal is presented for an intelligent, machine-learning-based adaptive learning framework, incorporating Attention-based Knowledge Tracing (AKT) to estimate continuous mastery, a calibrated gradient-boosted model to predict risk in early learning, and a contextual bandit policy to recommend adaptive learning content. The suggested closed-loop architecture dynamically examines learner interactions, such as correctness, time-on-task, and engagement indicators, to provide personalized and fair learning interventions. The framework is assessed using 1,200 students and 185,000 interaction records from a classroom-scale dataset. The experimental evidence shows better results than baseline models with an accuracy of 92, a smaller knowledge tracing error (RMSE = 0.17), a better calibration of the probability (ECE = 0.031) and much less fairness gap (F1-gap = 0.05). The results of these studies demonstrate that the suggested framework is effective in improving learning outcomes and minimizing learning differences, underscoring its applicability to online and blended learning.

Corresponding Author:

Hazim Noman Abed
Department of Computer Science, College of Science
University of Diyala
Baqubah City, Diyala Governorate, Iraq.
Email: hazim_numan@Uodiyala.edu.iq

1. INTRODUCTION

Learning gaps have also become chronic in many educational systems within different regions because of variations in preparation, teacher quality, learning resources, socio-economic factors and digital divide[1]. One pacing strategy of curriculum in many classes fails to deal with the differences in the abilities of the learners, thus leaving some learners behind, as others fail to be challenged[2]. These differences in learning eventually impact on academic results, self-esteem, motivation, and future career paths[3]. The recent trends in online learning and learning management systems (LMS) have presented the possibilities of adaptive learning data-driven [4]. Contemporary systems are able to record fine-grained learning records of quiz attempts, response times, the use of hints, frequency of revision, and patterns of engagement [5]. Such traces of digital learning allow computational algorithms to assess the mastery of students, identify single-subject misconceptions in their early stages, and prescribe intervention [6]. Nevertheless, it is difficult to construct dependable adaptive learning systems because of poor and noisy data of student behaviors, changing mastery trends with time and amplified risk of bias [7].

Classical methods in adaptive learning comprise classical methods like the Item Response Theory (IRT) and Bayesian Knowledge Tracing (BKT), and deep learning methods like Deep Knowledge [8]. Tracing (DKT). Attention based models like Self-Attentive Knowledge Tracing (SAKT) and Attention-based Knowledge Tracing (AKT) also enhance temporal modelling by recording long term interactions between interactions [9]. Parallel to this, reinforcement learning and multi-armed bandit strategies have been suggested to be used in adaptive sequence of the content. However, most of the studies are dedicated to either mastery prediction or policy optimization, but they do not construct an end-to-end framework [10].

1.1. Statement of the Problem

Although the field of digital education has been developed, most learning systems continue to use traditional and monotonic methods of instruction that do not meet the needs of different learners thus creating long-term learning gaps. Despite the fact that online platforms provide a vast amount of data about learner interaction, the current adaptive learning models tend to utilize this information rather inefficiently in terms of addressing disconnected processes, including but not limited to performance prediction or content recommendation instead of continuous mastery tracking, early signal detection, and adaptive decisions. Also, the noisy data problems, the lack of temporal modeling, inadequate calibration, and the problem of fairness are still to be addressed, and algorithmic personalization occasionally contributes to the existing imbalance. As a result, it is necessary to implement an intelligent, holistic, fairness-conscious adaptive learning system able to precisely capture the dynamics of knowledge, learners and its risks early on, and dynamically tailor content to enhance student performance in a fair manner.

1.2. Objectives of the Study

The current research proposes to design and test an adaptive learner's framework that is smart and efficient in filling learning gaps in education and optimizing student learning outcomes through the use of machine learning. The following are the specific objectives of the study:

1. To design an intelligent adaptive learning architecture that integrates student interaction data, knowledge tracing, and adaptive decision-making for personalized learning delivery.

2. To develop an accurate student mastery estimation model using attention-based machine learning techniques to capture the temporal evolution of learners' knowledge and skills.

3. To predict learning risk at an early stage by employing calibrated machine learning models capable of identifying students who are likely to fall behind academically.

4. To implement a dynamic content recommendation mechanism using contextual bandit–based adaptation strategies that personalize learning resources according to individual learner needs and engagement levels.

2. LITERATURE REVIEW

Sari et al. [11]explored how adaptive learning systems based on artificial intelligence can be used to enhance the educational outcomes by dynamically changing instructional material to suit individual learners. Their work tested the process of personalization by which learning routes were modified along performance, engagement, and interaction patterns of students. The results showed that AI-assisted adaptation was capable of contributing to the improvement of academic performance, efficiency, and motivation of learners significantly compared to traditional and non-adaptive models of instruction. Nevertheless, the research was mainly based on outcome-level improvements and system usability, not covering the elaboration of temporal knowledge tracing models and fairness and equity issues when it comes to adaptive decision makin.

Ezzaim et al. [12] presented a flexible learning model, which utilized a new machine learning-based performance forecasting model to enhance academic results. They used historical data of learners to forecast future learning and discover underachieving students at a very early age. Findings revealed that effective predictive modeling made possible timely interventions with instructional outcomes that resulted in quantifiable changes in the performance of learners as well as system responsiveness. The framework had these benefits though it did not incorporate continuous mastery evolution modeling and adaptive policy mechanisms to personalized content sequencing as its main aim .

Demartini et al. [13] provided a case study of how artificial intelligence can be used in adaptive learning in the real world. The experiment assessed the effects of AI-assisted personalization on student engagement, learning persistence, and academic achievement on the institutional level. The authors were able to state positive results in terms of higher participation of learners and better learning consistency which proved the feasibility of introducing AI into education as a practical one. Nonetheless, the case study method restricted the extrapolation of the results, and the design of the system did not have sophisticated temporal learning models, risk prediction calibration and a systematic assessment of fairness between groups of learners . Gligorea et al. [14] performed a substantial literature review of AI-based adaptive learning systems as applied to e-learning. Their literature review has systematically evaluated available methodologies and algorithms, as well as implementation issues, and concluded that artificial intelligence could greatly contribute to helping individuals to have personalized and scalable learning opportunities.

Notably, the authors have found that there are critical gaps in the research, such as the lack of explainable learning models, the lack of fairness and bias mitigation, and non-integrated end-to-end adoptive learning systems. Their study showed the necessity of smart structures that would integrate prediction, adaptation, and equity in an integrated system. Strielkowski et al. [15] explored AI-based adaptive learning as an instrument of sustainable learning change. Their analysis was centered on the long-term implications of adaptive learning technologies on the inclusiveness of the education systems, efficiency, and scalability.

The authors stated that AI-based personalization helped achieve better outcomes for learners and more efficient use of learning tools, thereby facilitating sustainability goals. The study, however, maintained a high level of strategic, conceptual rigor, with minimal information provided regarding the specifics of algorithmic implementation or the empirical validation of the mastery prediction and adaptive intervention strategies.

3. PROPOSED INTELLIGENT ADAPTIVE LEARNING FRAMEWORK

The growing complexity and diversity of contemporary learning systems require smart educational systems that can easily adapt to individual learners' needs in a timely and efficient manner. The conventional models of instruction that mostly use a predetermined sequence of content and a standardized way of assessment prove to be inadequate in meeting the differences in the prior knowledge, speed of learning and engagement of the learners. To address these shortcomings, this paper proposes a smart adaptive learning model that facilitates individualized and inclusive learning through machine learning approaches. The suggested framework is developed as a closed-loop, data-driven system that continuously records learner interactions, analyzes learning behavior, and dynamically adjusts interventions. The framework helps monitor learning progress in real time and anticipate learning problems by combining student profiling, feature engineering, mastery prediction, and adaptive decision-making within a single architecture. This holistic design not only contributes to reducing individual learning outcomes but also to reducing educational gaps, as it ensures that instructional support is responsive, equitable, and aligned with the changing needs of a particular learner.

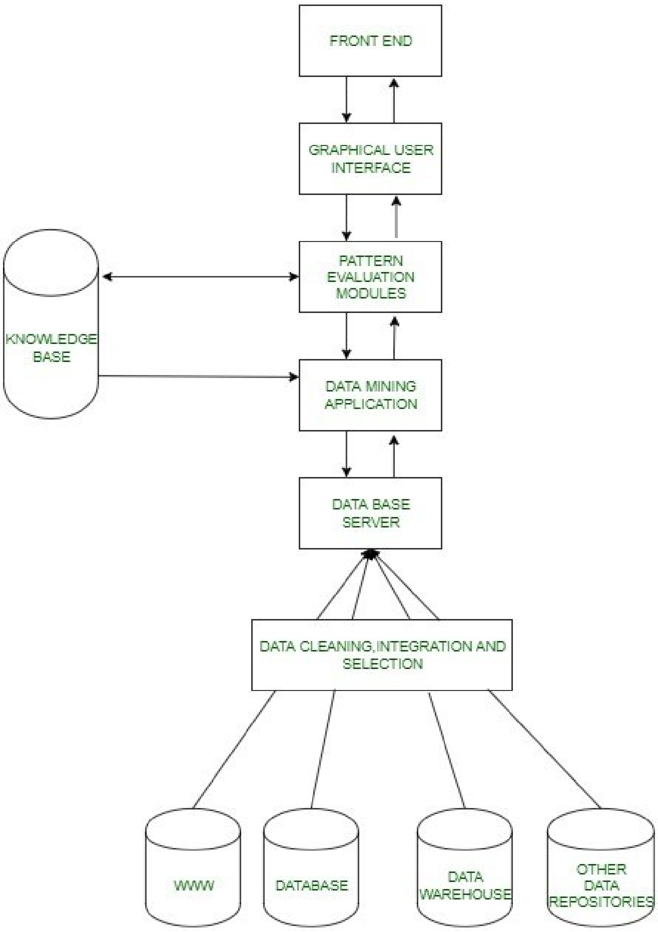

Figure 1. The block diagram of the data mining architecture shows the overall pipeline.

3.1. System Overview

The suggested intelligent adaptive learning model aims to provide both personalized and fair learning opportunities by continuously reviewing students' interactions and adjusting instructional plans accordingly. This framework also contrasts with a more traditional learning system that is considered to be a static system that involves a one-way flow of data, such as the learner, predictive modeling, and adaptive decision making, because it utilizes all three elements to support the various learning requirements in real time. This framework is designed in four large and interrelated modules, and each one plays a different yet complementary role:

1. **Student Data Acquisition and Profiling**: This module will gather finer learner interaction data in digital learning settings. It captures quiz attempts, response accuracy, response time, hints used, and content navigation behaviors and engagement indicators. Based on this data, a dynamic learner profile is generated

by the student, encompassing previous knowledge, learning speed, level of engagement, and favorable behavioral patterns.

2. **Feature Engineering and Preprocessing**: Raw interaction data tend to be sporadic, incomplete, and heterogeneous. The data is scrubbed and cleaned by this module using normalization, missing-value handling, and time sequencing. Significant characteristics are learned at different levels: interaction level, concept level, and student level to produce organized representations to be used in machine learning models. These are the characteristics that form the basis for proper knowledge and the anticipation of risks.

3. **Mastery Prediction and Risk Assessment**: Evaluation: Under this module, complex machine learning algorithms determine the changing level of mastery between students on various concepts. Attention-based knowledge tracing schemes are based on temporal patterns of learning, and a risk prediction model calibrates the likelihood that students will struggle and disengage. This twofold evaluation not only enables constant monitoring of learning progress but also enables the early detection of vulnerable learners.

4. **Adaptive Decision-Making and Intervention Policy**: Depending on the predicted mastery state and the risk level, this module determines the most suitable learning intervention that every student should have. A contextual decision-making strategy actively suggests individualized resources, e.g., remedial work, more complex problems, instructional videos, or feedback messages. The success of every intervention is measured by learner outcomes, creating a feedback loop that continually improves subsequent suggestions.

In general, the system operates in a continuous monitoring and feedback loop, in which learning recommendations are updated based on the student's progress. The framework would help enhance learning outcomes and minimize gaps in education by combining predictive intelligence with adaptive policies. Figure 1 shows the general flow of the proposed intelligent adaptive learning framework, emphasizing the sequential, feedback-based interactions among its main components. The diagram shows a closed-loop system in which learner data is continuously used to inform instructional decisions, enabling real-time personalization and improved learning outcomes. This pipeline starts with the student interaction layer, which discusses the student's interaction with digital learning materials such as quizzes, assignments, videos, and other learning materials. The interactions produce raw data on response correctness, response time, use of hints and indicators of engagement. This information is subsequently sent to the student data acquisition and profiling module, which consolidates and organizes learner data to build a dynamic student profile. Then, the received data is passed through the feature engineering and preprocessing unit, where missing values, noise, and inconsistencies are eliminated, meaningful features are extracted, and the unit performs feature engineering and preprocessing. The step converts raw interaction logs into structured representations that reflect learning behavior and time patterns with high accuracy, thereby enabling machine learning analysis.

The feature outcomes are then relayed to the mastery prediction and risk assessment module. In this case, attention-based knowledge tracing models can estimate learners' mastery levels for concepts, and the risk prediction component can identify students who may not be successful or are disengaged. Such predictive insights can offer an in-depth insight into the academic status of every learner. Depending on predictions of mastery level and risk, the adaptive decision-making and intervention policy module selects appropriate individualized learning actions. Such activities can include remedial support, challenging learning resources, educational videos, or targeted feedback. The learning platform delivers the selected interventions back to the learner.

3.2. Data Representation and Feature Engineering

Good data representation and feature engineering are essential to facilitating the proper modeling of student learning behavior. The activity of each learner in the proposed framework is represented as a time series of interactions, enabling the system to support both short- and long-term learning dynamics. The interaction sequence of a student sis is formally expressed as<|human|>. The interaction sequence of student sis is expressed as eq. (1):

$$D_s = \{(q_t, r_t, m_t)\}_{t=1}^{T} \ \ldots\ldots\ldots\ldots\ldots \ (1)$$

where $q_t$denotes the question or concept identifier attempted at time step t, $r_t$rerepresentsthe learner's response correctness encoded as a binary variable (1 for correct, 0 for incorrect), and $m_t$corresponds to auxiliary metadata associated with the interaction. The metadata also contains context related information such as time-on-task, the number of hints given, the type of device used, and the level of engagement that in the combination of the two factors will give a more insightful understanding of the learner behavior than just the correctness.

In order to appropriately leverage this interaction data, features are designed at three levels that are complementary to each other. Interaction level attributes record finer details about behavioral indicators in each learning event: response time, number of attempts and the use of hints, which are used to detect patterns such as hesitation, guessing or cognitive load. Skill-level features provide an overview of past performance by concept, including an estimated average mastery and an error rate per concept, and allow the system to monitor concept learning as well as potentially reveal long-term misconceptions. Student-level features are more general aspects of the learners, such as attendance history, pre-test performance ranges, and total engagement scores, which are used to derive the long-term trends in learning and individual variations. Considering that the educational data can be noisy and incomplete, the relevant preprocessing methods would be used to make it tough. Gaps in values are filled through median imputation of attributes which are not varying with time and temporal forward filling of attributes varying over time, such that learning sequences are continuous. Further, time-based features and engagement-related features are normalized to minimize the difference on the scales and enhance convergence of the model. This multi-level and systematic feature engineering process converts raw interaction logs into informative representations, which can be used to make trustworthy predictions of mastery and make adaptive learning choices.

3.3. Mastery Estimation Using Attention-based Knowledge Tracing (AKT)

To achieve successful personalization in adaptive systems, it is necessary to model the development of the knowledge of the student over the time. Attention-based Knowledge Tracing (AKT) is employed in the proposed framework to estimate the level of learners mastery by explicitly capturing dependencies across past learning interactions. In contrast to other recurrent methods like Deep Knowledge Tracing (DKT) which process interaction in a strictly sequentially manner and may loss lonk-range information, AKT uses self-attention mechanisms to learn long-range dependencies and assign adaptive importance weights of previous interactions events. This allows the model to focus on the most informative learning events such as repeated errors or consistent correct responses, while reducing the influence of less relevant past interactions. This allows AKT provides a more accurate and interpretable representation of the student knowledge development.

The latent student representation is calculated as at each time step t as eq (2):

$$h_t = AKT(h_{t-1}, q_t, r_t, m_t) \ldots\ldots\ldots\ldots \quad (2)$$

where h_t denotes the hidden knowledge state of the learner, $q_t$represents the question or concept attempted, $r_t$indicates response correctness, and $m_t$includes contextual metadata such as response time and engagement indicators. The attention mechanism aggregates information from previous interactions to form a context-aware representation of the learner's current knowledge state.

The probability of correctly answering the next item is then estimated as eq (3):

$$p_t = \sigma(Wh_t + b) \quad \ldots\ldots\ldots\ldots \quad (3)$$

Wand bare parameters are learnable and is the sigmoid activation function. The mastery vector M t obtained provides an idea of what the learner understands in real time across various concepts. This mastery estimation will be used as one of the inputs to the later risk prediction and adaptive decision-making modules to provide timely and specific instructional interventions.

3.4. Early Risk Prediction with Calibrated Gradient Boosting

It is important to identify students who are vulnerable to failing or losing interest in their studies early, so that adaptive learning settings can be effective. Within the framework proposed, the gradient-boosted decision tree model based on XGBoost is employed to pefrom early risk prediction using not only the generated latent knowledge representations by the AKT model, but also other performance-related attributes. The risk prediction model input is the mastery vector M_t, the current conceptual knowledge of the learner, and generated behavioral attributes f_t s such as recent performance patterns, variability in response time, and signs of engagement. The risk score at time step t= t is calculated as eq(4):

$$\rho_t = f_\phi(M_t, f_t) \ldots\ldots\ldots\ldots \quad (4)$$

$f_\phi(\cdot)$ (defined as the gradient-boosted model, parameterized by 0) (the 0 in this context is omitted). The output ρ t represents the likelihood of a student developing learning problems, poor performance or lack of interest in further learning activities. Probability calibration methods including temperature scaling or isotonic regression are used to increase the reliability and meaning of predictions. The methods are used to match the risk probability projections of the model with the observed failure rate of the actual value to ensure that the measurements of the model reflect the actual values of the real world. One area that risk estimates are most well-calibrated is in education where the adaptive systems and educators make informed decisions based on them to implement the right and proportional interventions.

Algorithm 1: Proposed Adaptive Learning using AKT + Risk Model + Bandit

Input: Student interaction stream D, resource set A, hyperparameters $\alpha, \beta, \gamma, \epsilon$

Output: Personalized adaptive action a_t

Initialize AKT parameters θ, risk model parameters ϕ, bandit parameters ψ
For each student $s$do
  Initialize mastery $M_0 = 0$
  For each interaction $t = 1..T$do
  Preprocess event (q_t, r_t, m_t)
    Compute hidden state $h_t = AKT_\theta(h_{t-1}, q_t, r_t, m_t)$
    Estimate mastery $M_t = \sigma(Wh_t + b)$
    Predict risk $\rho_t = GBM_\phi([M_t, f_t])$
    Create context $x_t = concat(M_t, \rho_t, engagement_t)$
    Select resource $a_t$using ContextualBandit(x_t, A, ϵ)
    Deliver $a_t$to student and observe reward R_t
    Update bandit parameters ψ
    Periodically update $\theta$and $\phi$with mini-batches
  End for
End for

4. EXPERIMENTAL SETUP

In order to prove the efficiency of the suggested adaptive learning framework that incorporates intelligence, a detailed experimental analysis was carried out based on a classroom-size dataset and in accordance with the commonly accepted educational benchmarking scales. The experimental design was meant to determine how well the framework can forecast student mastery and find learning threats, provide reliable estimates of probability, and foster equity among different classes of learners. The data in this paper was gathered in a real classroom-based digital learning setting and is provided in anonymized transcripts of student interaction. Through the comparison of the proposed method with the various conventional machine learning models and the existing knowledge tracing benchmarks, the evaluation will be a solid analysis of predictive performance, calibration quality, and equity results. This part will outline the dataset setup, metrics used in the evaluation, the baseline models and the quantitative findings of the results represented by Tables I-IV and corresponding graphical illustrations.

Table 1 describes the main characteristics of the dataset, as well as the experimental configuration. The dataset include 1,200 students of interaction information amounting and around 185,000 learning interactions records, covering 920 questions attributed to 110 skills or concepts. Each student completes an average 154 intreactions, which is long enough of interaction sequences to model learning behavior, providing appropriately long sequences to model temporal learning behavior. In every interaction record, there are the following attributes: student identifier (anonymized), question or concept ID, response correctness, response time, number of hints, and features related to engagement. The characteristics allow cognitive and behavioral modeling of the performance of learners. To ensure robust and unbiased evaluation, the dataset is split into training, validation, and testing sets using a 70/10/20 student-wise partition, preventing data leakage across subsets. The dataset is not publicly available because of institutional and privacy issues. Nevertheless, it has the same structure as popular educational data mining datasets like ASSISTments and EdNet, which increases the generalizability and reproducibility of the proposed structure.

Table 1. Dataset and Experimental Configuration

Parameter	Value
Students	1,200
Interactions	185,000
Questions	920
Skills/Concepts	110
Avg. Attempts per Student	154
Split	70/10/20 (student-wise)

The dataset structure can be characterized as realistic and large enough educational environment that could be used to test adaptive learning models. The range- of questions and skills permits mastery to be estimated with fine-grains at the concept level with the result of the large number of interactions allowing temporal knowledge to be accurately traced. The student separated analysis of the data avoids information leakage and represents the situation of real-world deployment, which enhances the validity of the experimental outcomes of the proposed framework. Table 2 gives a comparative analysis of the proposed adaptive learning framework with traditional machine learning and baseline models, evaluated in terms of predictive performance and calibration quality. Accuracy and AUC are used to evaluate the models in predicting mastery, next-question correctness prediction with RMSE, and Expected Calibration Error (ECE) to evaluate the probability reliability.

Table 2. Comparative Results (Prediction and Calibration Performance)

Model	Accuracy	AUC	RMSE ↓	ECE ↓
LR	0.78	0.82	0.25	0.062
SVM	0.81	0.84	0.23	0.057
RF	0.86	0.89	0.21	0.051
XGBoost	0.88	0.91	0.19	0.044
Proposed	0.92	0.94	0.17	0.031

The results demonstarte that the proposed framework outperforms all baseline models in all measures of evaluation. Specifically, it achieves the hieghest accuracy 0.92, and AUC 0.94, indicating superior predicative capability in modeling student mastery. Furthermore. lower RMSE (0.17) reflects imporoved accuracy in predicting subsequent reactions students, while smallest smallest ECE (0.031) proves that the model produces well-calibrated and reliable probability evaluations. These advancements underscore the efficiency of combining attention-based knowledge tracing and calibrated gradient boosting and allow the model to balance the dynamism of learning over time and the quality of risk estimation better than previous methods.

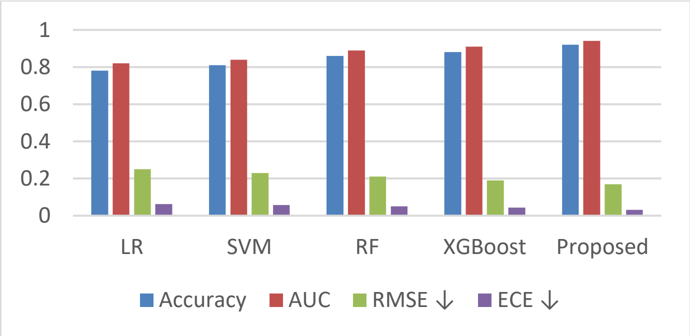

Figure 2. Comparative Performance of Models in Terms of Prediction Accuracy and Calibration Metrics)

Figure 2 illustrates the comparative performance of different models perform in terms of prediction and calibration metrices. It is evident from the figure that the proposed framework consistently outperforms the baseline models, achieving higher accuracy and AUC values while maintaining lower RMSE and ECE scores. These findings visually verify the numerical results of Table 2, and confirm that the proposed model is giving a high predictive performance, as well as, well-calibrated probability estimates. The fact that the graphical and tabular results are consistent also helps to support the robustness and stability of the suggested framework. Table 3 assesses the fairness of various models based on the F1-scores of high-performing and low-performing groups of students. The gap between these two scores is measured by the F1-gap measure, in which lower scores signify enhanced fairness and less unevenness in the performance of the groups.

Table 3. Fairness Evaluation Across Student Groups

Model	F1 (High-Performing Group)	F1 (Low-Performing Group)	F1-Gap ↓
LR	0.81	0.69	0.12
SVM	0.83	0.72	0.11
RF	0.87	0.78	0.09
XGBoost	0.89	0.81	0.08
Proposed Framework	0.92	0.87	0.05

The findings indicate that the proposed framework gives the lowest F1-gap (0.05), which implies that there is an excellent decrease in the gap in performance across various groups of students. This indicates that mastery estimation, the prediction of risks, and adaptive decision-making should be integrated to achieve more fair learning outcomes. Besides enhancing the general predictive accuracy, the framework suggested their effectiveness in ensuring fairness among all populations of learners, which is the most appropriate in real-world education settings where equity is a paramount concern.

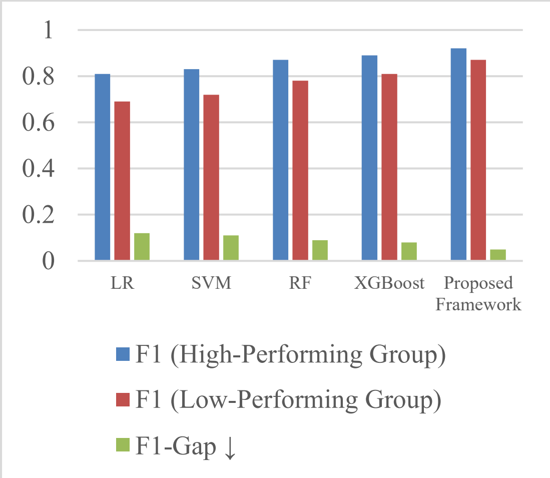

Figure 3. Fairness Comparison Across Models Based on F1-Scores for High- and Low-Performing Student Groups.

Figure 3 provides a visual representation of fairness performance across different models. The narrowed down performance gap between the high- and low-performing students of the proposed framework is also clearly evident, demonstrating a more balanced and equitable prediction capability through diverse learner groups. This decrease in inequality confirms the performance of the proposed fairness-conscious design and proves that the model can retain stable predictive accuracy and reduce bias between groups of students. This conduct is paramount to real-life education systems, which have fairness and inclusivity as the necessary features. Table 4 presents a comparison between the knowledge tracing abilities of the classical and deep learning-based model, using RMSE and average learning gain as evaluated metric. RMSE is the error of prediction and average learning gain determines the predictive capability of the model to help the students to maintain the learning process in the long run.

Table 4. Knowledge Tracing Performance Comparison

Knowledge Tracing Model	RMSE ↓	Average Learning Gain
BKT	0.24	0.41
DKT	0.21	0.47
SAKT	0.19	0.53
Proposed AKT-Based Model	0.17	0.61

The results show that the proposed AKT based model achives the lowest RMSE (0.17) and the highest average learning gain (0.61) as compared to the BKT, DKT, and SAKT models. This indicates the excellence of attention-related temporal modelling in not only long-term learning relationships but also in improving the mastery levels of students. The results show that the proposed AKT-based model achieves the lowest RMSE (0.17) and the highest average learning gain (0.61) among all compared methods. This implies high prediction and improved modeling of student knowledge development. These results indicate the efficiency of attention-driven temporal modeling, which allows the model to incorporate long-term dependencies in learning behavior. As a result, the suggested solution does not only enhance the level of prediction but also assists students in obtaining meaningful improvement in learning.

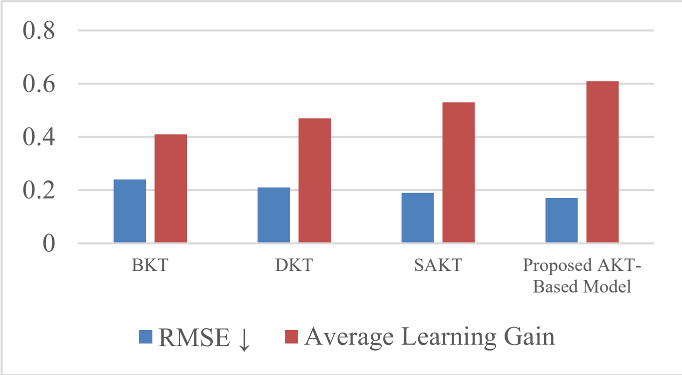

Figure 4. Knowledge Tracing Performance Comparison in Terms of RMSE and Average Learning Gain

Figure 4 compares the performance of various models of tracing knowledge, and it can be seen that there are variations in prediction error and learning gain. The suggested AKT-based approach has a clear advantage in comparison with baseline models, where the prediction error is lower, and the average learning gain is higher. This shows that the model is more useful in describing the dynamics of student learning. The graphical analysis is in line with the quantitative analysis, which was provided in Table 4, which once again demonstrates the reliability and strength of the suggested method. These findings support the idea that knowledge tracing based on attention offers important benefits in modeling long-term dependencies, resulting in better prediction ability and the support of student learning process.

In order to measure the strength and stability of the experimental results, the statistical validation was conducted on the results of the evaluation. In particular, paired statistics significance tests were performed to compare the proposed framework with baseline models based on the important performance measures such as accuracy, RMSE, and AUC. The paired t-test was used to determine the statistical significance of the improvements made. The outcome shows that the improvements of the proposed model are statistically significant with less than0.05 p-values signifying that the performance gains are not as a result of chance variation. Moreover, the computation of the evaluation metrics in terms of confidence intervals of 95% was done to determine the stability of the model performance. The low confidence levels of the suggested framework indicate similar behaviour of different subgroups of data. These statistical results are a good indicator that the proposed adaptive learning framework is not only better than the baseline models but also in a reliable and repeatable way.

5. CONCLUSION AND FUTURE WORK

In this paper, an intelligent adaptive learning scheme was described that overcomes some of the main challenges associated with personalized education such as the heterogeneous learning needs, the noisy nature of the interaction data, and the fairness issue. Combining knowledge tracking based on attention, risk prediction that is calibrated, and adaptive decision-making with the help of a contextual bandit, the proposed system will allow tracking the progress of learners continuously and providing personalized instructional interventions.

Experimental analysis of a classroom-sized dataset has shown that the framework is better than both traditional machine learning and available knowledge tracing methods, with greater accuracy in mastery prediction, smaller prediction error, and better calibration, as well as a smaller F1-gap between student groups. These findings confirm the efficiency of deep temporal modeling combined with calibrated prediction and adaptive policies as effective in improving the results of learning and suggesting equity. Altogether, the suggested framework provides a scalable and fairness-conscious solution to the contemporary educational platform, and future research will focus on large-scale real-world implementation, causal feedback on interventions, as well as the expansion of fairness bounds to a wider variety of learner features.

REFERENCES

[1] B. Singh and C. Kaunert, “Hidden Gems Breakthrough Dynamic Landscape of Adaptive Learning Technologies for Higher Education: Bridging the Gap Between Theoretical and Practical Knowledge Projecting Student Learning Outcomes,” in *Adaptive Learning Technologies for Higher Education*, IGI Global, 2024, pp. 222–247. DOI: 10.4018/979-8-3693-3641-0.ch010

[2] S. G. Essa, T. Celik, and N. E. Human-Hendricks, “Personalized adaptive learning technologies based on machine learning techniques to identify learning styles: A systematic literature review,” *IEEE Access*, vol. 11, pp. 48392–48409, 2023. DOI:10.1109/ACCESS.2023.3276439

[3] C. Halkiopoulos and E. Gkintoni, “Leveraging AI in e-learning: Personalized learning and adaptive assessment through cognitive neuropsychology—A systematic analysis,” *Electronics*, vol. 13, no. 18, p. 3762, 2024. DOI: 10.3390/electronics13183762

[4] C.-C. Lin, A. Y. Q. Huang, and O. H. T. Lu, “Artificial intelligence in intelligent tutoring systems toward sustainable

education: a systematic review," *Smart Learn. Environ.*, vol. 10, no. 1, p. 41, 2023. DOI: 10.1186/s40561-023-00260-y
[5] F. Naseer, M. N. Khan, M. Tahir, A. Addas, and S. M. H. Aejaz, "Integrating deep learning techniques for personalized learning pathways in higher education," *Heliyon*, vol. 10, no. 11, 2024. https://doi.org/10.1016/j.heliyon.2024.e32628.
[6] M. A. U. Zaman and E. Akhter, "Adaptive learning systems for English literature classrooms: A review of AI-integrated education platforms," *Int. J. Sci. Interdiscip. Res.*, vol. 4, no. 3, pp. 56–86, 2023. https://doi.org/10.63125/a30ehr12.
[7] T. Kabudi, I. Pappas, and D. H. Olsen, "AI-enabled adaptive learning systems: A systematic mapping of the literature," *Comput. Educ. Artif. Intell.*, vol. 2, p. 100017, 2021. DOI: 10.1016/j.caeai.2021.100017.
[8] S. Dutta, S. Ranjan, S. Mishra, V. Sharma, P. Hewage, and C. Iwendi, "Enhancing educational adaptability: A review and analysis of AI-driven adaptive learning platforms," 2024.
[9] A. Ezzaim, A. Dahbi, A. Aqqal, and A. Haidine, "AI-based learning style detection in adaptive learning systems: a systematic literature review," *J. Comput. Educ.*, vol. 12, no. 3, pp. 731–769, 2025. DOI:10.1007/s40692-024-00328-9.
[10] D. Hadar Shoval, "Artificial Intelligence in Higher Education: Bridging or Widening the Gap for Diverse Student Populations?," *Educ. Sci.*, vol. 15, no. 5, p. 637, 2025. DOI: 10.3390/educsci15050637.
[11] H. E. Sari, B. Tumanggor, and D. Efron, "Improving educational outcomes through adaptive learning systems using AI," *Int. Trans. Artif. Intell.*, vol. 3, no. 1, pp. 21–31, 2024. DOI:10.33050/italic.v3i1.647
[12] A. Ezzaim, A. Dahbi, A. Haidine, and A. Aqqal, "Enhancing academic outcomes through an adaptive learning framework utilizing a novel machine learning-based performance prediction method," *Data Metadata*, vol. 2, p. 164, 2023. DOI:10.56294/dm2023164.
[13] C. G. Demartini, L. Sciascia, A. Bosso, and F. Manuri, "Artificial intelligence bringing improvements to adaptive learning in education: A case study," *Sustainability*, vol. 16, no. 3, p. 1347, 2024. DOI: 10.3390/su16031347.
[14] I. Gligorea, M. Cioca, R. Oancea, A.-T. Gorski, H. Gorski, and P. Tudorache, "Adaptive learning using artificial intelligence in e-learning: A literature review," *Educ. Sci.*, vol. 13, no. 12, p. 1216, 2023. https://doi.org/10.3390/educsci13121216.
[15] W. Strielkowski, V. Grebennikova, A. Lisovskiy, G. Rakhimova, and T. Vasileva, "AI-driven adaptive learning for sustainable educational transformation," *Sustain. Dev.*, vol. 33, no. 2, pp. 1921–1947, 2025. DOI: 10.1002/sd.3221.

BIOGRAPHIES OF AUTHORS

Hazim Noman Abed is Associate Professor at College of Science, University of Diyala, Iraq. He received the B.Sc. degree in computer science from the University of Diyala and M.Sc. degree from UNITEN, Malaysia His research areas are Artificial Intelligence, cloud computing, and security. He has published several scientific papers in national, international conferences and journals. He can be contacted at email: Hazim_numan@uodiyala.edu.iq.