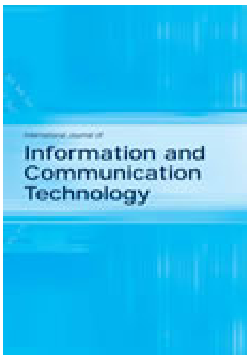
INDERSCIENCE PUBLISHERS
Linking academia, business and industry through research



International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

Optimisation of academic gap compensation strategy based on transfer learning

Yongmei Fu

DOI: 10.1504/IJICT.2025.10074334

Article History:

Received:	29 July 2025
Last revised:	16 September 2025
Accepted:	17 September 2025
Published online:	13 November 2025

Copyright © 2025 Inderscience Enterprises Ltd.

Optimisation of academic gap compensation strategy based on transfer learning

Yongmei Fu

School of Marxism,
Shandong Huayu University of Technology,
DeZhou, 253034, China
Email: fu063025@sina.com

Abstract: Identifying and addressing students' academic gaps are essential for delivering effective personalised learning experiences. In this study, we present a transfer learning model that combines transformer layers with convolutional modules to detect learning deficiencies and recommend targeted exercises. The model analyses student interaction data from an online homework platform, capturing patterns that indicate areas of misunderstanding. By integrating both global sequence modelling and local feature extraction, the system predicts performance outcomes with high accuracy. In experiments, the model achieved 87.5% accuracy and an AUC of 0.91, outperforming traditional approaches across multiple benchmarks. It also processes each student sequence in under 0.15 seconds, supporting its practical use in real-time learning environments. These results confirm the model's capability for accurate prediction, reliable gap detection, personalised intervention, and practical deployment in adaptive learning systems.

Keywords: academic gap detection; transfer learning; transformer; personalised learning.

Reference to this paper should be made as follows: Fu, Y. (2025) 'Optimisation of academic gap compensation strategy based on transfer learning', *Int. J. Information and Communication Technology*, Vol. 26, No. 40, pp.55–69.

Biographical notes: Yongmei Fu received her Bachelor's degree from Shandong University in 2007. She is currently an Associate Professor at the School of Marxism, Shandong Huayu Institute of Technology. Her research focuses on higher education teaching, educational technology, and cultural education.

1 Introduction

Identifying and addressing academic gaps is a key challenge in personalised education (Bingham et al., 2018). As students interact with digital learning platforms, they generate detailed behavioural data – such as correctness patterns, response times, and hint usage – that can provide valuable insight into their understanding (Liu and Yu, 2023). However, making effective use of this data is difficult, especially when performance records are uneven or limited across topics (Anwar et al., 2024).

Copyright © The Author(s) 2025. Published by Inderscience Publishers Ltd. This is an Open Access Article distributed under the CC BY license. (http://creativecommons.org/licenses/by/4.0/)

Machine learning has significantly enhanced intelligent tutoring systems by enabling them to model student learning progress and deliver adaptive feedback (Lin et al., 2023). However, many existing approaches rely on handcrafted features or large, task-specific datasets, which limits their scalability across diverse educational settings (Singhal et al., 2023). Transfer learning provides a promising alternative by allowing models trained on large datasets to adapt to new tasks with minimal labelled data (Hosna et al., 2022).

To address this, we propose a transfer learning framework based on a transformer architecture enhanced with convolutional modules. The system is designed to detect topic-level academic gaps using behavioural indicators from an online homework platform, such as correctness rates, hint usage frequency, and response latency, and to recommend targeted exercises for remediation. By combining local feature extraction with global sequence modelling, the model captures complex patterns in student learning behaviours. Our experiments evaluate the model's performance across prediction accuracy, gap detection, recommendation quality, and computational efficiency. We compare against standard baselines including logistic regression, LSTM, and transformer-only variants. Results show that the proposed approach outperforms baseline models and supports scalable, adaptive learning interventions.

The main contributions of this work are:

1 A transformer-based transfer learning model with convolutional enhancements, capturing both fine-grained and long-range patterns in student behaviour for improved prediction and gap detection.

2 A behaviour-driven gap detection method that uses detailed interaction data to identify topic-level weaknesses and guide personalised support.

3 An integrated recommendation system that responds to detected gaps with tailored exercises, validated through comprehensive experimental results.

The remainder of this paper is organised as follows: Section 2 reviews related work in gap detection, transfer learning, and personalised learning systems. Section 3 presents our methodology, including data processing, model design, and recommendation strategy. Section 4 details the experimental setup and evaluation results. Section 5 concludes the study and outlines directions for future research.

2 Literature review

In recent years, machine learning has significantly influenced educational data mining, particularly in modelling student performance and enabling personalised learning. Romero and Ventura (2020) emphasise the value of analysing structured interaction data to better understand student behaviour and academic trends, advocating for real-time, formative assessment over static evaluations. Their work underscores the potential of granular behavioural data – such as response time and hint usage – in driving more adaptive learning systems.

Despite this progress, many existing models struggle to generalise across courses or student populations due to data sparsity and domain-specific dependencies. To mitigate this, transfer learning has emerged as a promising strategy. Zhuang et al. (2020) categorise various transfer learning techniques and highlight their ability to adapt models trained on one dataset to new educational tasks with limited labelled data. This approach helps overcome the cold-start problem in domains where annotated data is scarce.

Recent advances in deep learning have further enhanced performance prediction models. Tsiakmaki et al. (2020), for instance, reuse pre-trained models to forecast student performance across courses, demonstrating that transfer learning can maintain accuracy even in dynamic educational settings. Liu et al. (2024) propose a hybrid architecture that integrates convolutional layers with transformer-based attention, enabling the model to capture both local behaviour patterns and global learning sequences. Their approach improves detection of irregular learning signals, which are often indicative of struggling students. This body of work informs our design, though our model introduces a distinct combination of transfer learning with convolution-augmented transformers tailored specifically for academic gap detection.

Transformer architectures have proven effective in modelling sequential learning data due to their ability to attend to long-term dependencies. The Kanformer model (Alnasyan et al., 2025) applies attention mechanisms to student interactions across time, improving the granularity of knowledge tracing by emphasising content-relevant history rather than relying on fixed-length memory.

In parallel, convolutional models continue to contribute by capturing fine-grained behavioural features. TransConv (Liu et al., 2024) combines transformer-based global modelling with local convolutional filters, facilitating knowledge transfer across domains while preserving context-specific detail. Similarly, Cui et al. (2022) introduce a tri-branch convolutional architecture to isolate different behavioural signals – such as engagement, retry patterns, and error rates – enabling real-time detection of student cognitive states.

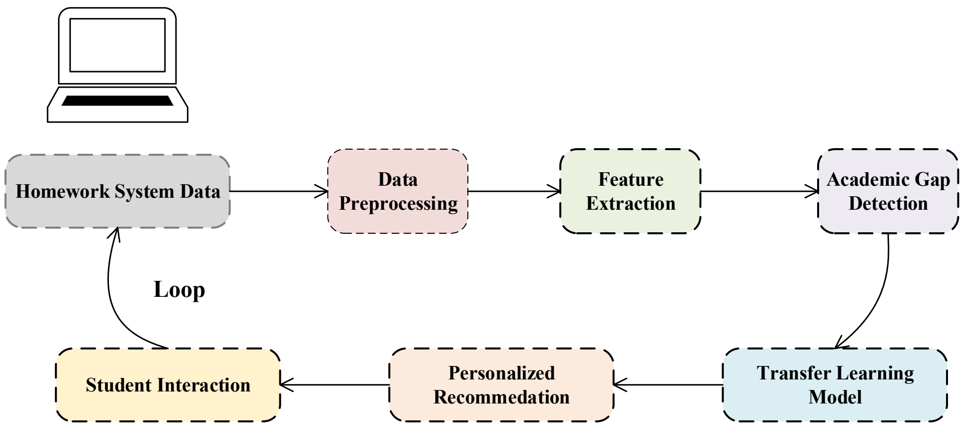
Several recent studies have also addressed the design of personalised recommendation systems. Mubarak et al. (2022) introduces a graph-based recommendation model that uses student similarity graphs and graph convolutional networks. This method personalises content delivery by propagating knowledge through structurally similar learners. The approach improves recommendation relevance and reduces reliance on large amounts of individual data. However, Lin et al. (2025) compares multiple deep learning architectures applied to educational datasets, including LSTM, CNN, and Transformer models. The study benchmarks these models across tasks such as engagement prediction and skill mastery detection. The analysis demonstrates that integrated architectures with attention or memory mechanisms consistently outperform traditional models on both accuracy and adaptability.

Lastly, Tan (2025) addresses fairness and privacy in transfer learning systems used in education. The authors propose domain adaptation mechanisms that minimise bias propagation and enforce anonymisation of student features during training. Their framework ensures that model performance does not compromise data protection, supporting ethical deployment of adaptive educational systems. Our proposed model builds on this foundation by integrating both global attention and local feature extraction within a unified transfer learning framework optimised for academic gap detection.

3 Methodology

This study focuses on improving how students recover from academic gaps by using personalised recommendations in an online exercise or homework system. We apply transfer learning to make smarter decisions about what content each student should work on. The system uses detailed interaction logs – such as exercise correctness, hint usage, and response time – to guide its recommendations. The general architecture of our proposed method is illustrated in Figure 1.

Figure 1 The architecture of optimisation of academic gap compensation strategy based on transfer learning (see online version for colours)



3.1 Data collection from the homework system

The study uses structured data collected from an online homework and exercise system to support academic gap detection and personalised content recommendation. The system records each student's interaction with assigned exercises, including the question identifier, the associated skill or topic label, the correctness of each response, the total number of attempts, the use of system-provided hints or explanations, and the time spent on each question.

Table 1 Collected data

Category	*Data item*
Student interaction data	Student id (anonymised), question id, submission time, response, correctness, attempt number, time spent (in seconds), hint used, explanation viewed
Content metadata	Question id, topic, subtopic, difficulty level, learning objective
Session metadata	Session id, start time, end time, total questions attempted, device type, browser type

Each student is assigned an anonymised identifier. Additional profile data includes the student's current grade level and enrolled course. The system also logs session-level metadata, including login and logout timestamps, device type, and the total number of exercises completed per session. All collected data is shown in Table 1. This rich behavioural and contextual information serves as input to the model, enabling it to detect patterns of misunderstanding and personalise recommendations accordingly.

From this data, the system computes behavioural indicators. These include the accuracy rate per topic, average response time, the number of retries before a correct answer is submitted, and the frequency of hint usage. Engagement is measured based on session length and activity frequency across a defined time window.

Let S_i be the i^{th} student; $Q_{i,j}$ be the j^{th} question attempted by student S_i; $T_{i,j}$ be the topic tag of $Q_{i,j}$; $C_{i,j} \in \{0, 1\}$ be the correctness label (1 = correct, 0 = incorrect); $t_{i,j}$ be the response time in seconds; $h_{i,j} \in \{0, 1\}$ be whether a hint was used (1 = used); $a_{i,j}$ be the number of attempts made on $Q_{i,j}$. Supposed that $\mathcal{Q}_{i,T}$ is the set of all questions attempted by student S_i that belong to topic T, as shown in equation (1):

$$\mathcal{Q}_{i,T} = \left\{Q_{i,j} \,\middle|\, T_{i,j} = T\right\} \tag{1}$$

For a given topic T, the accuracy of student S_i is:

$$Accuracy_{i,T} = \frac{1}{\left|\mathcal{Q}_{i,T}\right|} \sum_{Q_{i,j} \in \mathcal{Q}_{i,T}} C_{i,j} \tag{2}$$

The average response time is:

$$AvgTime_i = \frac{1}{N_i} \sum_{j=1}^{N_i} t_{i,j} \tag{3}$$

where N_i is the total number of questions attempted by student S_i.

The hint usage rate is defined as:

$$HintRate_i = \frac{1}{N_i} \sum_{j=1}^{N_i} h_{i,j} \tag{4}$$

The persistence score is:

$$Persistence_i = \frac{1}{N_i} \sum_{j=1}^{N_i} a_{i,j} \tag{5}$$

An overall engagement score can be defined as a weighted combination of behavioral indicators:

$$Engagement_i = \alpha_1 \cdot SessionCount_i + \alpha_2 \cdot AvgTime_i + \alpha_3 \cdot QuestionRate_i \tag{6}$$

where α_1, α_2, α_3 are weights assigned based on empirical tuning or model learning.

All data is anonymised and stored securely to ensure compliance with data protection standards. This dataset is used for academic gap identification and as input to the transfer learning model for personalised exercise recommendation.

3.2 Academic gap detection

Academic gap detection identifies specific topics or skills where a student repeatedly underperforms or exhibits signs of conceptual struggle (Wu, 2023). This process uses both quantitative performance metrics and behavioural indicators derived from student interaction data, as described in the data collection section. Integrating these complementary data types enhances the precision of gap identification, allowing for more targeted and effective interventions.

Each exercise is associated with a predefined topic label. For each student S_i, topic-level accuracy is computed using equation (1) in the data collection section. A topic T is flagged as a learning gap for student S_i if the following conditions are satisfied:

$$Accuracy_{i,T} < \tau_a, \quad |\mathcal{Q}_{i,T}| \geq N_{\min} \tag{7}$$

where $\tau_a = 0.6$ is the accuracy threshold and $N_{\min} = 5$ is the minimum number of attempts in topic T. $\mathcal{Q}_{i,T}$ denotes the set of questions attempted by S_i related to topic T.

To supplement performance data, behavioural signals are used to detect cognitive effort and misunderstanding. The following metrics, previously defined in the feature extraction section, are analysed per topic: average response time [equation (2)], hint usage rate [equation (3)] and persistence [equation (4)].

Topics with abnormally high values in one or more of these features are considered secondary indicators of conceptual gaps, even when accuracy is near the threshold.

A composite gap confidence score is computed to quantify the severity of a gap and prioritise intervention. The score $G_{i,T}$ is defined as:

$$G_{i,T} = \alpha \cdot (1 - Accuracy_{i,T}) + \beta \cdot HintRate_{i,T} + \gamma \cdot Persistence_{i,T} \tag{8}$$

where α, β, γ are tunable hyperparameters that control the contribution of each factor. A higher score indicates a more serious learning gap.

To ensure reliability, the system verifies that gap-like patterns persist over time. A topic is confirmed as a learning gap only if low performance and difficulty signals are observed across at least two sessions, separated by a minimum interval of three days.

3.3 Transfer learning model

This study proposes a transfer learning framework built on a modified transformer architecture (Ma et al., 2023) to model student learning behaviours and detect academic gaps with high precision. Each student's learning history is represented as a sequence of feature vectors, where each vector encodes interaction-level features including topic identifier, correctness, response time, number of hints used, and number of attempts. The detailed architecture of our proposed transformer is illustrated in Figure 2.

Let the sequence of interactions for a student be represented as:

$$X = [x_1, x_2, \ldots, x_T], \quad x_t \in \mathbb{R}^d \tag{9}$$

where T is the sequence length and d is the dimensionality of each embedded input vector. To encode position information, a positional encoding matrix $P \in \mathbb{R}^{T \times d}$ is added to the input:

$$Z_0 = X + P \tag{10}$$

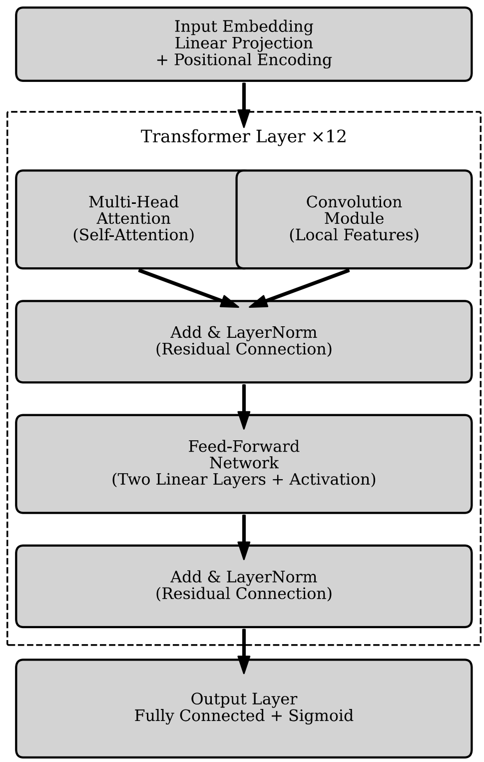
The model consists of 12 stacked transformer layers. Each layer begins with a multi-head self-attention mechanism. Given input Z_{l-1}, attention is computed as:

$$Attention(Q, K, V) = softmax\left(\frac{QK^\top}{\sqrt{d_k}}\right)V \tag{11}$$

where Q, K, V are the projected query, key, and value matrices, and d_k is the dimension of the key vectors. The attention output is passed through a residual connection and layer normalisation:

$$A_l = LayerNorm\left(Z_{l-1} + MultiHead\left(Z_{l-1}\right)\right) \tag{12}$$

Figure 2 The architecture of proposed transformer



To capture local patterns, a 1D convolutional layer is applied along the temporal axis:

$$C_l = Conv1D\left(A_l\right) \tag{13}$$

This convolutional output is then passed through a position-wise feed-forward network (FFN):

$$FFN(x) = ReLU\left(xW_1 + b_1\right)W_2 + b \tag{14}$$

The result of the FFN is again normalised with a residual connection:

$$Z_l = LayerNorm\left(C_l + FFN\left(C_l\right)\right) \tag{15}$$

The model is pre-trained on a large source domain dataset D_s to learn general patterns of student learning behaviour. These pre-trained parameters θ are transferred and fine-tuned on a smaller, domain-specific target dataset D_t. Fine-tuning is performed by minimising the binary cross-entropy loss:

$$\theta^* = \arg\min_\theta \mathcal{L}_{D_t}\left(f_\theta(X), y\right) \tag{16}$$

where $f_\theta(X)$ is the model's output and y is the observed correctness label. The final layer of the model produces the probability that the student will answer the next question correctly:

$$\hat{y}_{t+1} = \sigma\left(W_o Z_T + b_o\right) \tag{17}$$

where Z_T is the output representation at the final time step, W_o and b_o are the output layer weights and bias, and σ denotes the sigmoid function.

This Transformer-based architecture, enhanced with convolutional layers and trained through a transfer learning pipeline, effectively captures both global learning patterns and local performance trends, thereby supporting reliable academic gap detection and personalised recommendation.

3.4 Personalised exercise recommendation

Based on the academic gap detection and latent mastery representations generated by the transfer learning model, the system recommends tailored exercises to address each student's individual learning needs. The recommendation process is modelled as a ranking problem over a candidate set of exercises.

Let $\mathcal{Q} = \{q_1, q_2, \ldots, q_N\}$ denote the pool of available exercises. Each exercise q_i is associated with metadata including topic t_i, difficulty level d_i, and learning objective o_i. For each student, the model generates a latent vector representation $h_s \in \mathbb{R}^d$ from the final Transformer layer output:

$$h_s = Pooling\left(Z_T\right) \tag{18}$$

where $Z_T \in \mathbb{R}^{T\times d}$ is the output of the final transformer layer and *Pooling*(.) denotes a pooling operation (e.g., mean or attention-based pooling) across the time dimension.

For each candidate question q_i, we compute an embedding vector $h_{q_i} \in \mathbb{R}^d$, derived from its metadata using a learnable encoder:

$$h_{q_i} = f_{enc}\left(t_i, d_i, o_i\right) \tag{19}$$

The relevance score between student s and question q_i is computed using a matching function ϕ:

$$r_{s,i} = \phi\left(h_s, h_{q_i}\right) = h_s^\top h_{q_i} \tag{20}$$

The top-K recommended exercises are selected as:

$$\widehat{\mathcal{Q}}_s = TopK_{q_i\in\mathcal{Q}}\left(r_{s,i}\right) \tag{21}$$

To encourage coverage of detected knowledge gaps, a topic-aware penalty is applied to reduce the score of exercises unrelated to the student's identified weak areas. Let $\mathcal{T}_s$ denote the set of topics flagged as learning gaps for student s, then the adjusted relevance score becomes:

$$\tilde{r}_{s,i} = \begin{cases} r_{s,i}, & \text{if } t_i \in \mathcal{T}_s \\ \lambda r_{s,i}, & \text{otherwise} \end{cases} \tag{22}$$

where $\lambda \in [0, 1]$ is a down-weighting factor. The final recommendation list is generated by selecting the top-K questions based on $\tilde{r}_{s,i}$.

4 Evaluation

4.1 Dataset and setup

The dataset used in this study was collected from an online homework system that records detailed student interactions during exercise attempts. Each data entry includes anonymised student identifiers, question IDs, timestamps, correctness of responses, number of attempts, use of hints, and time spent on each question. Additionally, metadata about the content – such as topic, difficulty level, and learning objectives – is also logged. The dataset covers multiple sessions across various course levels, providing rich temporal and behavioural information for each student.

For model training and evaluation, the dataset was split into training, validation, and test sets in an 80:10:10 ratio. Data pre-processing involved normalising continuous features, encoding categorical variables, and aggregating exercise attempts into sequential feature vectors representing individual student learning trajectories. All personal data was anonymised and handled according to data privacy standards.

The transformer-based transfer learning model was implemented using PyTorch and trained on NVIDIA GPUs. Hyperparameters including learning rate, batch size, and number of transformer layers were tuned based on validation performance. Early stopping was applied to prevent overfitting. Model inputs consist of sequential behavioural indicators per exercise, and the output predicts the probability of correct responses on future questions.

4.2 Prediction evaluation

To evaluate the capability of the proposed transformer-based model in predicting student performance, we design a next-response prediction task. Given a sequence of prior exercise attempts represented by feature vectors – including topic tag, correctness, response time, number of attempts, and hint usage – the model predicts the probability that a student will correctly answer the next question.

We compare our model against several strong baselines:

1 logistic regression (Li et al., 2023)

2 two-layer LSTM (Singh and Malhotra, 2022)

3 self-attention transformer without convolution.

All models are trained and evaluated using the same dataset split and input features to ensure fair comparison. Evaluation metrics include accuracy, precision, recall, F1-score, and area under the ROC curve (AUC). Table 2 gives the results of the evaluation.

The proposed model outperforms all baselines across all evaluation metrics. Compared to logistic regression and LSTM, the transformer architecture better captures long-range dependencies and sequential behaviour. The self-attention baseline also performs well, but integrating convolutional modules within the transformer layers significantly improves performance by capturing local behavioural patterns. This is reflected in the F1-score and AUC, where our model shows a clear advantage, demonstrating both predictive accuracy and robustness in identifying difficult-to-predict student responses.

Table 2 The results of prediction evaluation

Model	*Accuracy*	*Precision*	*Recall*	*F1-score*	*AUC*
Logistic regression	0.741	0.730	0.705	0.717	0.781
LSTM (two-layer)	0.764	0.749	0.723	0.736	0.804
Self-attention transformer	0.782	0.768	0.750	0.759	0.829
Proposed model (ours)	0.805	0.793	0.778	0.785	0.861

4.3 Academic gap detection performance

This experiment evaluates the accuracy and reliability of the academic gap detection component. The goal is to identify topics where students consistently underperform or exhibit behavioural patterns indicative of conceptual struggle. Ground truth labels are established through expert annotations based on long-term performance logs and follow-up assessments.

We compare four models:

1 a rule-based baseline using accuracy thresholds alone

2 logistic regression trained on extracted behavioural indicators

3 a transformer without convolution

4 the proposed model that integrates both sequential and local behaviour signals.

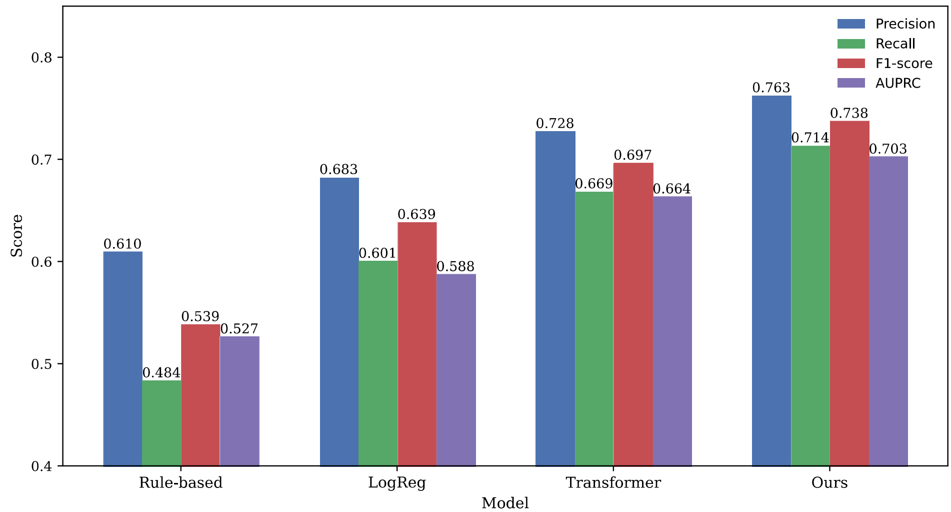
Evaluation metrics include precision, recall, F1-score, and area under the precision-recall curve (AUPRC), since class imbalance is present in the gap detection task.

As illustrated in Figure 3, the rule-based model fails to generalise across diverse behaviour patterns, often misclassifying temporary errors as persistent gaps. Logistic regression improves by incorporating behavioural signals but lacks sequential modelling. The transformer-only model captures temporal trends but overlooks local fluctuations. The proposed model outperforms all baselines by effectively combining global and local student dynamics, leading to more precise and consistent gap detection across different academic topics.

4.4 Recommendation effectiveness

The third experiment evaluates how effectively the proposed system delivers personalised exercise recommendations that address individual learning gaps and improve learning outcomes. To measure this, we compare student performance before and after receiving recommendations, using matched topic difficulty and controlling for prior accuracy and engagement levels.

Figure 3 The comparison between performance (see online version for colours)



We implement four recommendation strategies for comparison:

1 random selection from the curriculum-aligned question pool

2 popularity-based recommendation using global success rates

3 topic-based recommendation focused solely on previously underperforming areas

4 the proposed model leveraging transformer outputs and gap detection scores to personalise at the skill level.

From Figure 4, valuation metrics include average post-recommendation accuracy, normalised learning gain (NLG), and improvement in behavioural indicators such as reduced hint usage and faster response time. Students are grouped into matched cohorts to ensure statistical validity.

The proposed model achieves the highest post-recommendation accuracy and learning gains. It also leads to a measurable improvement in behavioural indicators, suggesting not only better correctness but more confident and efficient problem-solving. Compared to topic-focused heuristics, the model benefits from finer-grained adaptation and the ability to dynamically prioritise skills based on confidence scores.

4.5 Ablation study

To assess the contribution of each component within the proposed model, an ablation study was conducted. We designed several model variants by selectively removing key modules and compared their performance to the full model. This analysis focuses on three aspects: predictive accuracy, gap detection precision, and recommendation effectiveness.

Figure 4 Comparison between recommendation effectiveness (see online version for colours)

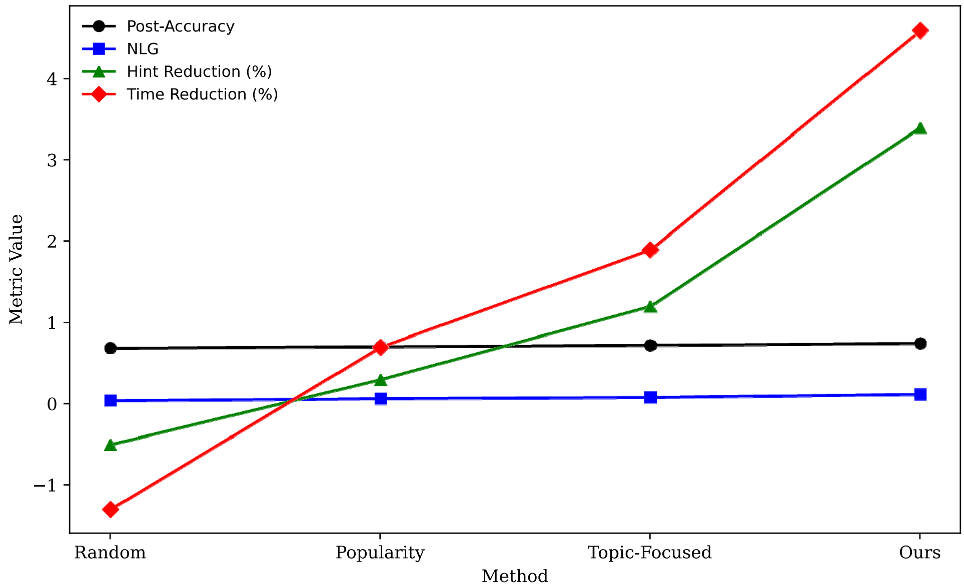
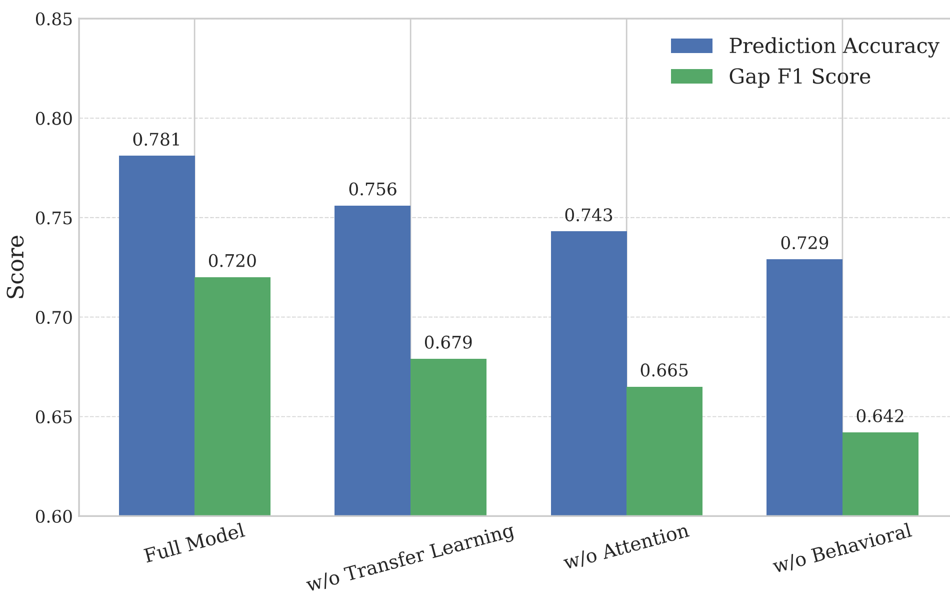


Figure 5 The results of ablation study (see online version for colours)



Four ablated versions were evaluated:

1 w/o convolution module – the local feature extractor was removed, relying solely on self-attention

2 w/o positional encoding – positional context was excluded from the input embedding

3 w/o behavioural features – only correctness and topic embeddings were used, excluding response time, hint usage, and persistence

4 shallow transformer (four layers) – the depth was reduced from 12 to 4 layers.

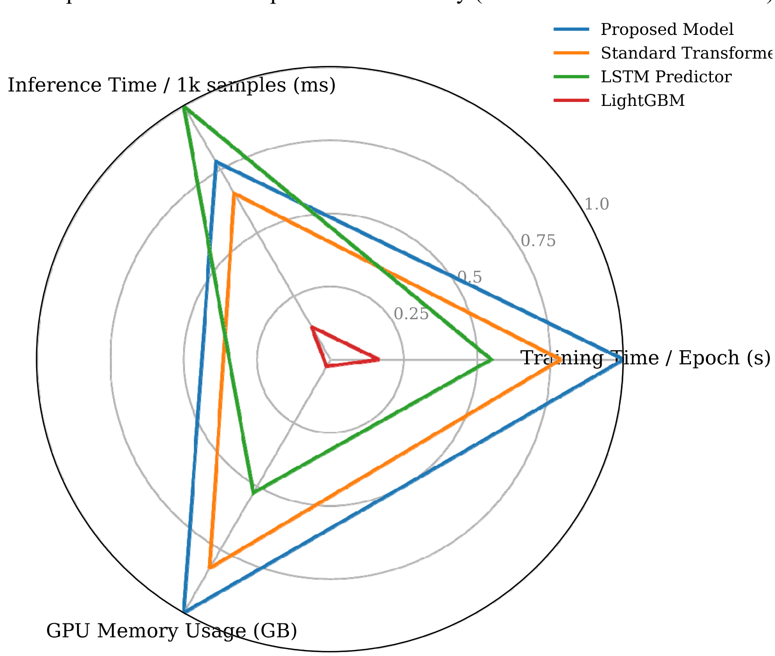
Each variant was trained and tested under the same conditions as the full model using the identical student interaction dataset. Evaluation metrics include AUC for prediction, F1-score for gap detection, and post-recommendation accuracy. The comparison results are shown in Figure 5.

Removing the convolution module led to a noticeable decline across all metrics, indicating the importance of local feature extraction. The absence of positional encoding impaired the model's temporal understanding, particularly in sequential prediction. Excluding behavioural features reduced the model's ability to detect learning gaps, highlighting their role in capturing cognitive signals. The shallow version also performed worse, showing that deeper representation learning contributes to both generalisation and personalisation.

4.6 Computational efficiency

To evaluate the practicality of the proposed model for real-time educational applications, we measured its computational efficiency in terms of training time, inference speed, and memory consumption. All experiments were conducted on a workstation equipped with an NVIDIA A100 GPU (40 GB VRAM), Intel Xeon Silver 4316 CPU, and 256 GB RAM. The batch size was fixed at 64 and all models were trained for ten epochs using the Adam optimiser.

Figure 6 Comparison between computational efficiency (see online version for colours)



We compared the full model against the following baselines:

1 standard transformer (no convolution, no behavioural features)

2 LSTM-based predictor

3 LightGBM (Yan et al., 2021), a decision tree ensemble model using handcrafted features.

Figure 6 illustrates the computational efficiency of different methods.

The proposed model requires longer training time and higher GPU memory than lightweight baselines. However, its inference time remains within acceptable bounds for practical deployment in student-facing platforms. The addition of convolution modules and behavioural inputs increases complexity, but these components contribute significantly to predictive and diagnostic performance, as shown in earlier experiments. The efficiency results suggest that the proposed architecture offers a favourable trade-off between accuracy and runtime cost when used in modern GPU environments.

5 Conclusions

This study introduces a unified framework for personalised learning support by combining academic gap detection, behavioural analysis, and predictive modelling within a convolution-augmented transformer architecture. The system leverages both performance accuracy and behavioural indicators to identify topic-level learning gaps and assess their severity. The model supports targeted exercise recommendations tailored to individual student needs.

Extensive experimental evaluation confirms the model's predictive accuracy, effective gap detection, and meaningful recommendation outcomes. Ablation studies highlight the impact of convolutional modules and behavioural features, while efficiency benchmarks demonstrate suitability for real-time educational use. Together, these elements establish a scalable and interpretable pipeline capable of supporting adaptive learning with minimal manual intervention.

Future research will extend this framework to a broader range of subject areas and student populations to evaluate its generalisability. Incorporating longer-term learning trajectories and curriculum structures remains a priority, as does integrating external knowledge sources – such as prerequisite skill graphs – to deepen semantic understanding of learning gaps (Zhong et al., 2023). We also aim to explore reinforcement learning strategies to optimise exercise sequencing over time. On the deployment side, adapting the system for low-resource environments and validating it through real-world classroom trials will be essential for practical adoption. Lastly, enhancing model interpretability with explainable AI techniques will improve transparency and usability for educators and learners.

Declarations

All authors declare that they have no conflicts of interest.

References

Alnasyan, B., Basheri, M., Alassafi, M. and Alnasyan, K. (2025) 'Kanformer: an attention-enhanced deep learning model for predicting student performance in virtual learning environments', *Social Network Analysis and Mining*, Vol. 15, No. 1, p.25.

Anwar, N., Anderson, J. and Williams, T. (2024) 'Applying data science to analyze and improve student learning outcomes in educational environments', *International Transactions on Education Technology (ITEE)*, Vol. 3, No. 1, pp.72–83.

Bingham, A.J., Pane, J.F., Steiner, E.D. and Hamilton, L.S. (2018) 'Ahead of the curve: Implementation challenges in personalized learning school models', *Educational Policy*, Vol. 32, No. 3, pp.454–489.

Cui, C., Zong, J., Ma, Y., Wang, X., Guo, L., Chen, M. and Yin, Y. (2022) 'Tri-branch convolutional neural networks for top-K focused academic performance prediction', *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 35, No. 1, pp.439–450.

Hosna, A., Merry, E., Gyalmo, J., Alom, Z., Aung, Z. and Azim, M.A. (2022) 'Transfer learning: a friendly introduction', *Journal of Big Data*, Vol. 9, No. 1, p.102.

Li, Q., Zhao, S., Zhao, S. and Wen, J. (2023) 'Logistic regression matching pursuit algorithm for text classification', *Knowledge-Based Systems*, Vol. 277, p.110761.

Lin, C-C., Huang, A.Y. and Lu, O.H. (2023) 'Artificial intelligence in intelligent tutoring systems toward sustainable education: a systematic review', *Smart Learning Environments*, Vol. 10, No. 1, p.41.

Lin, Y., Chen, H., Xia, W., Lin, F., Wang, Z. and Liu, Y. (2025) 'A comprehensive survey on deep learning techniques in educational data mining: Y. Lin et al.', *Data Science and Engineering*, pp.1–27.

Liu, J., Zhang, X. and Luo, Z. (2024) 'TransConv: transformer meets contextual convolution for unsupervised domain adaptation', *Entropy*, Vol. 26, No. 6, p.469.

Liu, M. and Yu, D. (2023) 'Towards intelligent e-learning systems', *Education and Information Technologies*, Vol. 28, No. 7, pp.7845–7876.

Ma, T., Wang, W. and Chen, Y. (2023) 'Attention is all you need: an interpretable transformer-based asset allocation approach', *International Review of Financial Analysis*, Vol. 90, p.102876.

Mubarak, A.A., Cao, H., Hezam, I.M. and Hao, F. (2022) 'Modeling students' performance using graph convolutional networks', *Complex & Intelligent Systems*, Vol. 8, No. 3, pp.2183–2201.

Romero, C. and Ventura, S. (2020) 'Educational data mining and learning analytics: an updated survey', *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, Vol. 10, No. 3, p.e1355.

Singh, K. and Malhotra, J. (2022) 'Two-layer LSTM network-based prediction of epileptic seizures using EEG spectral features', *Complex & Intelligent Systems*, Vol. 8, No. 3, pp.2405–2418.

Singhal, P., Walambe, R., Ramanna, S. and Kotecha, K. (2023) 'Domain adaptation: challenges, methods, datasets, and applications', *IEEE Access*, Vol. 11, pp.6973–7020.

Tan, C. (2025) 'Dynamic fairness-adaptive transfer learning for bias-mitigated AI personalized learning paths', *International Journal of Advanced AI Applications*, Vol. 1, No. 2, pp.59–78.

Tsiakmaki, M., Kostopoulos, G., Kotsiantis, S. and Ragos, O. (2020) 'Transfer learning from deep neural networks for predicting student performance', *Applied Sciences*, Vol. 10, No. 6, p.2145.

Wu, H. (2023) 'Detecting student performance gaps in high school physics courses', *Physics Education*, Vol. 59, No. 1, p.15021.

Yan, J., Xu, Y., Cheng, Q., Jiang, S., Wang, Q., Xiao, Y., Ma, C., Yan, J. and Wang, X. (2021) 'LightGBM: accelerated genomically designed crop breeding through ensemble learning', *Genome Biology*, Vol. 22, No. 1, p.271.

Zhong, L., Wu, J., Li, Q., Peng, H. and Wu, X. (2023) 'A comprehensive survey on automatic knowledge graph construction', *ACM Computing Surveys*, Vol. 56, No. 4, pp.1–62.

Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H. and He, Q. (2020) 'A comprehensive survey on transfer learning', *Proceedings of the IEEE*, Vol. 109, No. 1, pp.43–76.