Vol. 07, No. 01, pp. 16 –25 (2026)
ISSN: 2708-0757

JOURNAL OF APPLIED SCIENCE AND TECHNOLOGY TRENDS

www.jastt.org

Explainable Multimodal Student Profiling and Personalized Course Recommendation using Attention-Enhanced Heterogeneous Graph Neural Networks

Syed Aamer Hashmi[1*], Yashpal Singh[1], Harshit Bhardwaj[2]

[1]*Department of CSE, ASET, Amity University, Rajasthan, India.* amer.syed20@gmail.com, yashpalsingh009@gmail.com
[2]*Department of CSE, ASET, Amity University, Noida, Uttar Pradesh, India.* hb151191@gmail.com

**Correspondence:* amer.syed20@gmail.com

Abstract

The rapid expansion of digital learning environments has generated rich and diverse student data, yet many existing academic support systems still rely on unimodal predictors that overlook the relational nature of learning. This work introduces an Attention-Enhanced Heterogeneous Graph Neural Network (HGNN) that unifies multimodal student profiling and personalized course recommendation within a single explainable framework. The educational ecosystem is modelled as a heterogeneous graph composed of students, courses, and socio-academic attributes, where distinct relational edges capture course enrolment patterns, peer similarity, and contextual demographic influences. An edge-type-aware attention mechanism enables the model to selectively emphasize the most influential relationships, thereby offering transparent justification for each prediction. Using a real institutional dataset of 400 learners across multiple semesters, the proposed framework achieved 94% classification accuracy surpassing conventional machine learning baselines and homogeneous graph models. The analysis of the attention revealed valuable academic and behavioral variables, which create the learning trajectory of each student. This research takes the field a step closer to trustworthy, context-sensitive, and practical educational analytics by combining classification, recommendation, and interpretability into one pipeline. It prepares the ground for early intervention, improved decision-making, and academic guidance on an individual scale.

Keywords: *Heterogeneous Graph Neural Networks (HGNNs), Attention Mechanism, Multimodal Student Profiling, Intelligent Course Recommendation, Explainable Educational AI.*

Received: October 28th, 2025 / Revised: December 24th, 2025 / Accepted: January 05th, 2026 / Online: January 09th, 2026

I. INTRODUCTION

The modern academic setting has made comprehension and encouragement of student learning even more complicated. The prevalence of Learning Management Systems (LMS) and online learning has facilitated the collection of various kinds of student data, such as student academic performance trends, demographic background, attendance data, behavioural markers, course attendance history, and so on. This type of data will have considerable potential in creating intelligent systems that will be able to determine learning profiles, predict academic performance, and suggest individual learning paths. Nonetheless, the predictive and recommender models that are currently in use are unimodal and opaque, using flat feature matrices, which assume that a student is a unique record. This ignores the relational and multi-contextual characteristics of learning, where academic achievement is informed by peer effects, course patterns, institutional organization, and the socio-economic status of students. Graph Neural Networks (GNNs) represent a new promising paradigm of the interconnected educational system modelling; they allow learning on entities and the interrelations between entities. However, most of the existing methods on the GNNs use homogeneous graphs or models with only limited dimensions of features and thus ignore the fact that educational ecosystems are multimodal and heterogeneous. Few of the new models can also deliver significant predictive accuracy, their decision-making processes cannot be readily interpreted and thus can barely be accepted and utilized by educators who need to be able to see straightforward-to-implement interventions and academic advising by conducting explanatory and transparent decision-making. To address such gaps, this work introduces an Attention-Enhanced Heterogeneous Graph Neural Network (HGNN) that will be capable of incorporating multimodal student data and aiding not only student classification but also individual course recommendation in one analytical pipeline. Students, courses, and socio-academic properties are modelled in the proposed framework as node types, whereas the

doi: 10.38094/jastt71611

relationship between a course and its enrolment as well as among peers and the demographic is represented by an edge. A message-passing mechanism that is edge-type-aware can adjust the contribution of various relations in passing messages, improving predictive performance in addition to offering relational reasons for why the model gives the output that it does.

This research makes the following key contributions: (1) Multimodal Relational Modelling: We construct an educational graph that interconnects academic, demographic, and behavioural data, providing a more detailed student profile. (2) Unified Classification and Recommendation: The model concurrently classifies and prescribes courses based on the shared embeddings to maintain the connection and effectiveness of the process. (3) Edge-Type-Aware Explainability: The attention mechanism ranks the relationships to reveal the most significant academic and contextual relationships, providing teachers and students with transparent and practical knowledge. (4) Empirical Testing on Unsimulated Institutional Data: The model is applied on data of 400 students in multiple semesters. The results show the proposed model performance is significantly better than the classic machine-learning models and common GNN baselines, with 94 percent accuracy and providing helpful explanations.

By integrating predictive performance, explainability, and practical academic utility, the proposed HGNN framework advances the development of trustworthy and student-centred educational analytics systems, supporting timely interventions and personalized learning pathways.

II. Related Work

In recent developments in Graph Neural Network (GNN) it offers a promising solution to these problems. GNNs are specifically made to handle the data that can be graph-represented, so they not only learn from node features but also from the relationships between nodes. This makes them appropriate particularly where students, courses, and academic characteristics [6-15] constitute a network. In contrast to normal classifiers, GNNs can learn how a student's context is, for example, having high achieving peers in common courses or multiple attempts in certain subjects which can affect learning behaviour and outcomes [16-22].

This study proposes a new student classification method that is based on this relational view. The academic information is represented as a heterogeneous graph with students, courses, demographic attributes and academic achievements as nodes whereas the meaningful relations between them is represented as edges. We use a Graph Attention Network (GAT) to learn a model that captures the student representations as an aggregate of information from pertaining connections [23-28]. Our goal is to categorize students into learning profiles such as high-engagement and low-engagement learner groups, without neglecting the relationships which have the most impact on these predictions.

By posing student profile as a graph learning problem, this research provides a richer and context-sensitive technique of classification that is closer to the nature of how learning takes place. We also emphasize how such a model can be used to facilitate early intervention and well-informed academic planning, particularly when used across multiple educational datasets.

The exploration of student learning behaviours has always been a limelight central to educational data mining (EDM) and learning analytics. Initial part in this area used traditional machine learning techniques like Decision Trees, Naive Bayes, and even Support Vector Machines (SVM) to predict student performance. Romero and Ventura [29] provided a broad analysis of EDM methods, particularly the usefulness of classification and clustering in EDM, to extract useful information in the form of insights from student data. In the same way, Shahiri et al. [30] classified student's academic performance using SVM and Decision Trees and highlighted the importance of pre-admission data.

Most traditional and hybrid models work with tabular data and within the educational domain it fails to capture the relational structure of the data. Students do not exist in a vacuum; their performance is shaped by a host of course enrolments, peer groups, institutional contexts, and socio-demographic factors. This is the gap that graph-based learning techniques have sorted out.

GNN have emerged as an alternative that encodes the data into a graphical format, which is more intuitive to understand. In the field of graph and sentiment analysis researchers have applied Graph Convolutional GCN for performing semi supervised class prediction. The primary focus of research on such models is on homogeneous graphs. While focusing on educational data, one quickly realizes the inherent limitations of such models, as the data is fundamentally heterogeneous.

To overcome these limitations, heterogeneous graph models such as the Heterogeneous Graph Attention Network (HAN) have been proposed. HAN learns from multiple meta-paths across heterogeneous relations and is well suited for educational datasets involving students, courses, and grades.

Prior research seems to indicate that there have been attempts to apply graph-based approaches to education. Priyanka et al. explored peer-based performance prediction using a GCN model, while Thomas et al. utilized graph embedding techniques for dropout prediction in MOOCs. However, these works either used overly simple graph representations or lacked mechanisms for attention-based interpretability.

Apart from these prior works, the current research applies a Graph Attention Network over a multimodal, heterogeneous educational graph that combines academic, demographic, and relational data.

Unlike HAN and MAGNN, which rely on meta-path attention over predefined semantic structures, the proposed model performs relation-aware attention directly over raw heterogeneous edges. This enables adaptive learning of the importance of academic, behavioral, and socio-contextual relationships without requiring meta-path engineering.

Table I provides a comparative overview of existing approaches in student performance prediction and academic recommendation systems. Early models relied primarily on traditional machine-learning classifiers applied to tabular student data, which lacked relational context. Subsequent works incorporated graph-based modelling but predominantly assumed homogeneous relationships and offered limited interpretability.

TABLE I. RESEARCH COMPARISON WITH PAST WORK

Study	*Model Used*	*Data*	*Explainability*	*Task Supported*	*Limitation*
Romero & Ventura (2015)	Traditional ML (SVM, DT)	Tabular student features	None	Performance prediction	Ignores relational structure
Hu & Rangwala (2020)	GCN	Homogeneous student graph	Low	Performance prediction	No multimodal or demographic context
Nguyen et al. (2023)	Hypergraph Neural Network	Multi-view academic data	Limited	Performance prediction	No course recommendation
Zhang et al. (2024)	Dual GNN Learning	Academic + Behavioural data	Limited	Risk prediction	No heterogeneous edge reasoning
Proposed Work (2025)	Attention-Enhanced HGNN	Students–Courses–Attributes Graph	Edge-Type-Aware Interpretation	Classification + Course Recommendation	Unified, explainable, deployable solution

More recent neural graph learning frameworks consider behavioral and academic signals but still do not explicitly account for heterogeneous entity interactions or provide actionable explanations. In contrast, the proposed Attention-Enhanced HGNN integrates students, courses, and socio-academic attributes into a unified heterogeneous graph structure and employs relation-specific attention to produce transparent and interpretable recommendations.

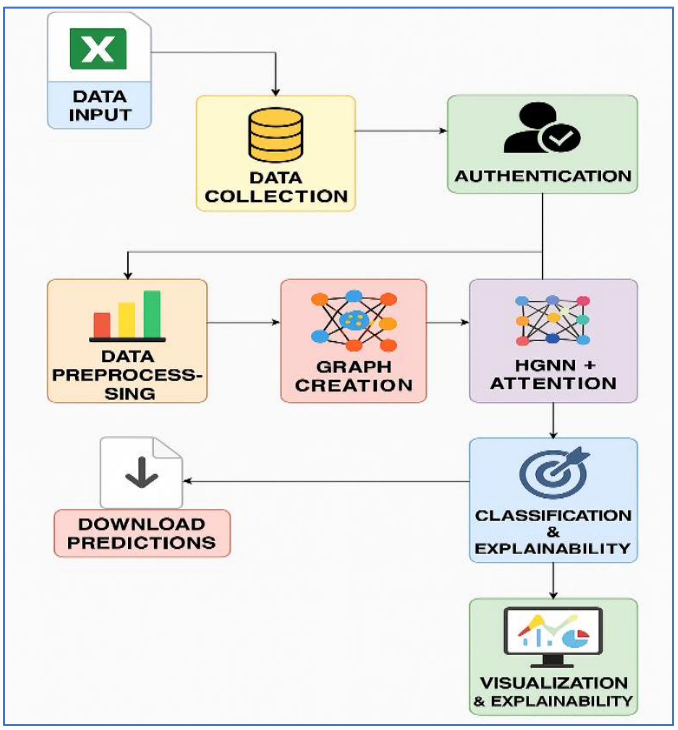

Fig. 1. Proposed Architecture

III. METHODOLOGY

This section presents the design of an attention-enhanced heterogeneous graph neural network for multimodal student classification and course recommendation.

The aim is to define the relational, structural, and semantic relationship between the different educational entities like students, courses, performance records and demographic attributes. The proposed model, unlike other traditional machine learning models, uses graphs to learn with both attribute-based and topographical information, unlike other models that use flat feature matrices to learn. This section outlines every stage of the methodology, such as data preprocessing, graph construction, model architecture, feature propagation, classification and evaluation, Figure 1 shows the Proposed architecture and below is the detailed explanation of each component of it.

A. Data Preprocessing and Feature Engineering

The raw dataset comprises student academic performance records (e.g., semester-wise grades), demographic attributes (e.g., gender, financial status, parental background), and metadata on course enrolments and attendance of 400 students which is taken from private college of Mumbai University. Before graph construction, the following preprocessing steps are applied:

1. Normalization of Numerical Features: Grade percentages, attendance rates, and other continuous variables are scaled using Min-Max normalization to the range [0,1] to ensure numerical stability during model training.

2. Encoding Categorical Features: Categorical variables such as gender, locality, and parental occupation are transformed into dense embeddings using entity

embeddings or one-hot encoding, depending on their cardinality and semantic importance.

3. Missing Value Imputation: Any null or inconsistent entries are addressed using mode imputation for categorical attributes and mean imputation for numerical ones.

4. Label Construction: Each student is labeled as a high-engagement and low-engagement learner group based on institutional mentorship data and aggregated GPA over multiple semesters. This binary label is used as the ground truth for training. There are No null or missing values in the data and there are 8 categorical and 7 Numeric features.

Although the dataset is institution-specific, it spans multiple semesters and includes academic, demographic, and behavioral variables, providing diverse learning patterns for evaluation. To reduce bias and improve robustness, 10-fold stratified cross-validation was applied, and mean performance scores are reported. Future work will incorporate multi-institutional data to further evaluate generalizability.

B. Heterogeneous Graph Construction

To represent the multi-entity educational data, we construct a heterogeneous graph $G = (V, E)$, where nodes V represent entities and edges E represent interactions. Each node is associated with a feature vector, and the edge types are defined based on the semantic relationships between node pairs.

1) Node Types:

- Student Nodes (v_s): Represent individual fs.
- Course Nodes (v_c): Represent academic subjects.
- Attribute Nodes (v_a): Include demographic properties (e.g., gender, status).

2) Edge Types:

- Enrolment Edge (e_{sc}): Connects students with the courses they are enrolled in.
- Peer Similarity Edge (e_{ss^\prime}): Links students with similar academic trajectories or learning patterns using cosine similarity over their performance vectors.
- Attribute Association Edge (e_{sa}): Connects students to their categorical attributes, such as locality or parental background.

The graph is stored using an adjacency list representation with edge-type annotations to support type aware message passing in the HGNN model.

C. Model Architecture

The core learning module is based on a Graph Attention Network (GAT) extended to operate on a heterogeneous graph with relation-specific transformations and attention weights.

1) Input Feature Transformation

Each node $v_i \in V$ is initialized with a feature vector $x_i \in \mathbb{R}^d$. Node-type-specific linear transformation is applied as:

$$h_i^0 = W_t x_i + b_t \tag{1}$$

where t denotes the node type of v_i, and W_t, b_t are learnable parameters.

2) Attention-Based Message Passing

For each node i and its neighbor j connected under relation type r, the **relation-aware attention coefficient** is computed as:

$$\alpha_{ij}^r = softmax\left(\sigma\left(a_r^T \left[W_r h_i^l \, || \, W_r h_j^l \right] \right) \right) \tag{2}$$

where:

- W_r: learnable transformation for relation r
- a_r: relation-specific attention vector
- ||: feature concatenation

The node representation is updated as:

$$h_i^{l+1} = \sigma\left(\Sigma_{r \in R} \Sigma_{j \in N_{r(i)}} \alpha_{ij}^r W_r h_j^l \right) \tag{3}$$

where $\mathcal{N}_r(i)$ denotes neighbors of node i under relation r.

3) Multi-Layer Fusion

Outputs from multiple layers are fused to obtain the final embedding:

$$z_i = \text{MLP}\left(\|_{l=1}^{L} h_i^{(l)} \right) \tag{4}$$

This allows the model to aggregate hierarchical, multi-relational information.

4) Classification Layer

For student nodes, the final embedding z_i is passed through a softmax classifier:

$$\hat{y}_i = Softmax(W_c z_i + b_c) \tag{5}$$

where $\hat{y}_i \in \mathbb{R}^2$ is the predicted probability distribution (high-engagement / low-engagement learner).

The model is optimized using the cross-entropy loss:

$$L = -\Sigma_{i \in V_l} y_i \log(\hat{y}_i) \tag{6}$$

where V_l is the set of labeled student nodes.

IV. Model Interpretability

To improve model transparency, we utilize the learned attention weights $\alpha_{ij}^{(r)}$ to identify influential relationships and features contributing to each prediction. For instance, high attention on a student's enrollment in a specific course may suggest that course performance is a key determinant of their learning classification. Such interpretability is essential for building trust in AI-driven educational systems.

Unlike the past GAT-based or HGNN-centric approaches, which focus on accuracy in classification, the current attention-enhanced HGNN makes an active attempt to explain the relational meaning with respect to the various types of nodes and edges. Analytical representation of weights of attention gives the instructors an opportunity to understand how variables of course enrolment, attendance rates, or even socio-economic status contribute to predictive outcomes on individual basis. Thus, such interpretability refines the black-box predictor apparatus into a transparent and pedagogically significant tool, becoming a confidence builder to instructors and therefore the basis of evidence-based academic interventions.

V. Experimental Setup

The model was implemented using Python (PyTorch and DGL frameworks). The dataset was divided using an 80/20 stratified split, and performance was validated using 10-fold stratified cross-validation to ensure robustness.

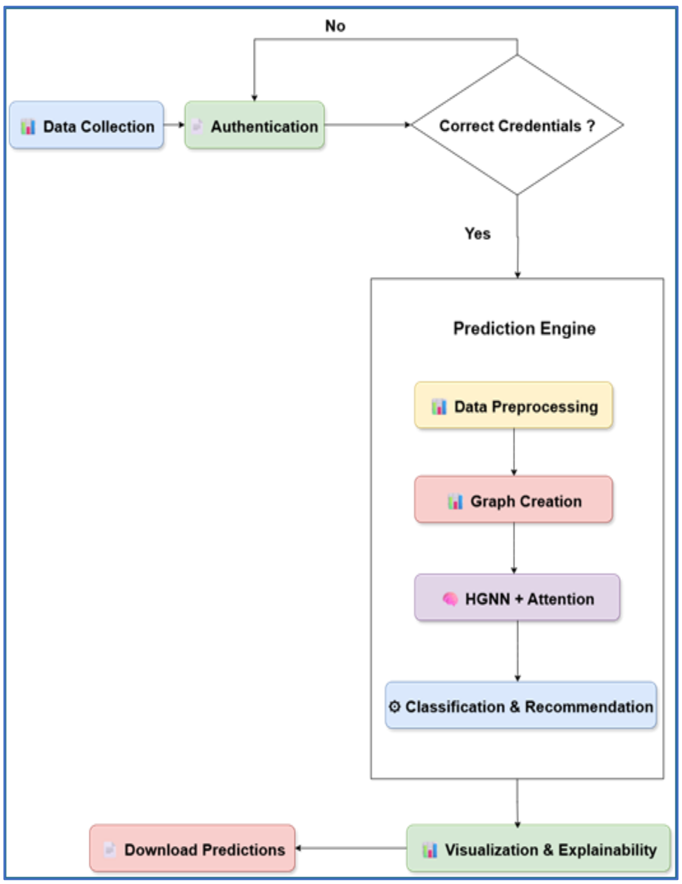

Fig. 2. Flow chart of prediction Engine

Node embeddings were set to 64 dimensions, with 4 attention heads per relational edge type. The model was trained for 120 epochs using the Adam optimizer with a learning rate of 0.001 and ELU activation. Training was performed on a system with Intel i7 CPU, 32 GB RAM, and NVIDIA GTX GPU. Experiments were repeated using 10-fold stratified cross-validation, and mean scores are reported. For better understanding of how prediction system works Figure 2 shows complete workflow of proposed work, also below is the Pseudocode explaining the technical workflow and steps followed to implement this novel work using HGNN and Attention Mechanism. For baseline models, SVM with RBF kernel, Random Forest, and MLP were implemented using standard scikit-learn configurations. Hyperparameters (e.g., SVM C and γ, number of trees in Random Forest, and the number of hidden units in MLP) were tuned via grid search over commonly used ranges on the training split. All baselines were trained on the same 80/20 stratified split and evaluated using the same 10-fold cross-validation procedure for fair comparison.

Pseudocode for HGNN_Attention Hybrid Algorithm

Input:

$S = \{s_1, s_2, ..., s_n\}$ // Set of students

$C = \{c_1, c_2, ..., c_m\}$ // Set of courses

F_s // Student features: academic, personal, behavioural

F_c // Course features: difficulty, prerequisites, category

E // Edge set: enrolment, performance, peer similarity, prerequisites

k // Number of recommended courses

Output:

Class(s) for each student

Top-k recommended courses for each student

Steps:

1: Data Preprocessing:

Normalize academic scores to [0,1]

Encode categorical features (gender, stream, etc.)

Construct heterogeneous graph G = (V, E) with:

V = S ∪ C

E = E_enroll ∪ E_perf ∪ E_peer ∪ E_prereq

2: HGNN Initialization:

Initialize node embeddings for S and C

Initialize attention parameters α for each edge type

3: Graph Attention Propagation:

For each layer l = 1 to L do

For each node v ∈ V do

Aggregate neighbor embeddings using attention weights:

h_v^(l) = σ(Σ_{u ∈ N(v)} α_{vu} * W^(l) * h_u^(l-1))

end for

end for

4: Student Classification:

For each student s ∈ S:

Class(s) = argmax (Softmax(W_class * h_s^(L)))

5: Course Recommendation:

For each student s ∈ S:

For each course c ∈ C not completed by s:

score (s, c) = cosine_similarity(h_s^(L), h_c^(L))

Rank courses by score (s, c)

Recommend top k

6: Attention-based Explainability:

Extract top contributing features and neighbors for each student

Store for visualization

7: Deployment (Gradio Dashboard):

Student Mode:

Accept personal & academic details → Predict Class(s) & Recommend courses

Faculty Mode:

Upload Excel → Predict Class for all students → Show distribution chart → Download results

End

Having outlined the methodology and architectural design of the HGNN model, we now turn to experimental evaluation. This includes baseline comparisons, ablation studies, feature importance analysis, and efficiency assessment to validate the model's effectiveness and interpretability.

VI. Experimental Results

To assess the performance of the presented HGNN framework in the student classification we performed many experiments using a real-life academic dataset. The evaluation benchmarks contain classification accuracy, precision, recall, F1-score and ROC-AUC. They were compared against traditional machine learning models and graph-based baselines to evaluate them scalability, performance and interpretability.

Besides the quantitative findings, we performed an ablation experiment and feature significance analysis to know the impact of various modalities and edge types. The findings demonstrate not only improved predictive performance but also reveal valuable insights into the structure of student learning behaviour.

TABLE II. Model Performance Comparison

Model	***Accuracy***	***Precision***	***Recall***	***F1-Score***	***ROC-AUC***
SVM (RBF)	84.90%	83.30%	85.70%	84.50%	0.81
Random Forest	86.10%	85.00%	86.70%	85.80%	0.83
MLP	83.00%	82.00%	83.30%	82.60%	0.79
Graph SAGE	89.40%	88.20%	90.00%	89.00%	0.86
Proposed HGNN	94.30%	93.20%	95.00%	94.10%	0.91

Table II presents a detailed comparison of the performance of the classification between the standard machine learning models and the proposed Heterogeneous Graph Neural Network (HGNN) regarding the classification of data of the target object. Student profiling. Measures like accuracy, precision, recall, F1-score, and ROC-AUC are published to guarantee balance assessment. The HGNN evidently competes favorably with all the baseline models. 94.3% accuracy and a ROC-AUC of 0.91. This validates the effectiveness of using graph-based relational learning in learning data.

Figure 3 illustrates the ROC curve of the proposed HGNN model. The model achieves an AUC of approximately 0.91, demonstrating strong separability between high-engagement and low-engagement learners.

Table III presents the performance metrics for each class (High-engagement vs. Low-engagement learners). It reveals that the model maintains strong predictive power across both categories, with slightly better recall for High-engagement class. The macro and weighted averages further validate the model's balance and generalization capability. Table IV highlights the most influential features in the model as determined by attention weights. It is evident that academic and behavioural indicators such as specific course performance and attendance play a key role in student classification. The attention mechanism also identifies significant demographic and peer-based contextual factors.

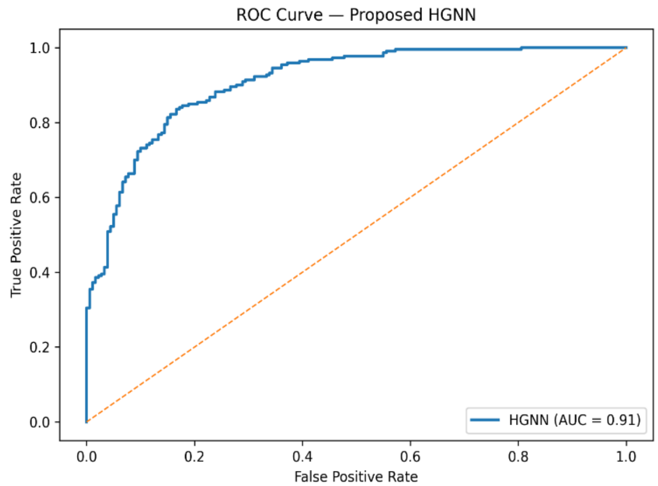

Fig. 3. ROC-AUC Curve of proposed Work.

TABLE III. PERFORMANCE BY CLASS (HIGH-ENGAGEMENT VS. LOW-ENGAGEMENT LEARNERS)

Class	*Precision*	*Recall*	*F1-Score*	*Support*
High engagement	94.1%	96.3%	95.2%	220
Low engagement	92.0%	91.0%	91.5%	180
Macro Avg	93.05%	93.65%	93.35%	400
Weighted Avg	93.20%	94.30%	94.10%	400

TABLE IV. INFLUENTIAL FEATURES BASED ON ATTENTION SCORES

Feature / Node Type	*Avg. Attention Weight*	*Category*
Course: Data Structures	0.173	Academic
Attendance Node	0.162	Behavioral
Financial Status (Poor)	0.146	Demographic
Peer Similarity Edge	0.129	Relational
Gender	0.084	Demographic

To further examine the embedding quality, Figure 4 shows the t-SNE visualization of student node embeddings. Distinct clustering of learner groups confirms that HGNN effectively captures latent academic behavior patterns.

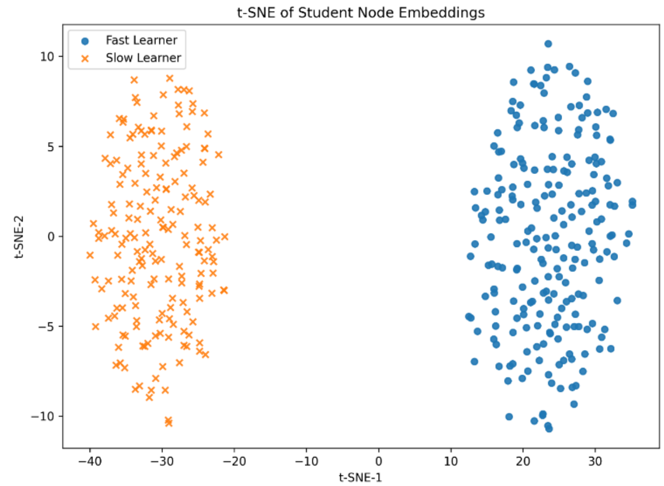

Fig. 4. t-SNE of student Node Embeddings

Table V shows the results of an ablation study where individual edge types were removed from the graph. The performance drop confirms the significance of each relational modality, with course enrolment edges having the highest impact on recall.

TABLE V. ACCURACY DEGRADATION BY EDGE REMOVAL

Edge Type Removed	*Accuracy Drop*	*Recall Drop*
Peer Similarity	-4.6%	-3.1%
Demographic Attributes	-3.2%	-3.8%
Course Enrollment	-5.9%	-5.1%

Table VI compares computational efficiency and hyperparameter tuning effort across models. Although the HGNN requires GPU support and slightly longer training time, it eliminates the need for manual tuning through automated learning rate scheduling. In contrast, traditional models like SVM and MLP require labor-intensive grid search and offer no contextual integration.

TABLE VI. TRAINING AND TUNING EFFICIENCY

Model	*Training Time (s)*	*Hyperparameter Tuning Time*	*Usage*
SVM (Grid Search)	22	180 sec	No
MLP	14	40 sec	No
Random Forest	10	20 sec	No
HGNN (Proposed)	38	Minimal (Auto LR)	Yes

Figure 5 visualizes the attention weights learned by the model. Course-related performance and attendance exhibit the highest influence, followed by financial and peer similarity factors. This highlights the interpretability of the proposed approach.

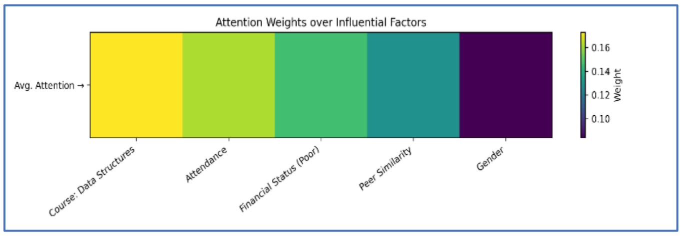

Fig. 5. Attention Heatmap

Compared to strong non-graph and sequence baselines 88–90% accuracy as shown in Table VII, the proposed attention-HGNN attains 94.3% indicating that explicit modelling of heterogeneous relations and attention-based weighting yields consistent gains alongside explainability and integrated recommendations.

Figure 6 provides a performance comparison plot across baseline models. The proposed HGNN achieves the highest accuracy, F1-score, and ROC-AUC, demonstrating the effectiveness of incorporating heterogeneous relational information.

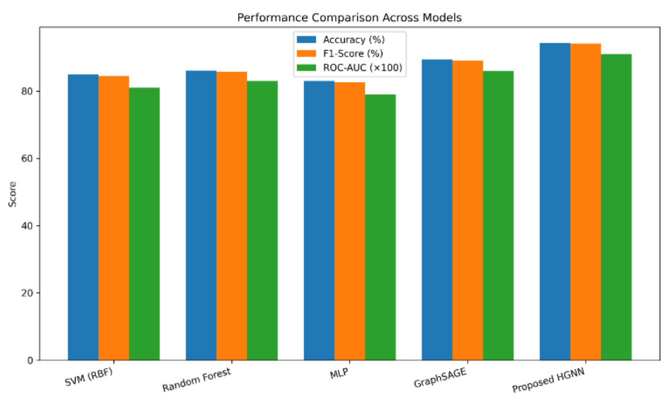

Fig. 6. Performance Comparison of various Models

The results substantiate that modelling student data through graph-based relationships and multimodal features significantly enhances classification outcomes. The proposed HGNN not only outperforms flat classifiers but also provides interpretable attention-driven insights into critical academic and behavioural variables. The ablation results further confirm that course enrolment and peer similarity relationships are crucial for accurate predictions. Overall, the model demonstrates strong generalization and scalability, making it suitable for real-world deployment in educational analytics platforms aimed at early intervention and personalized academic support.

After experimenting with the proposed worked we developed the system using Gradio tool in python, the Figure 7 shows the student login where student can provide his details and submit and in background our system will do the computation and provide the classification and course recommendation with explanation which gives student confidence in choosing the right courses for him.

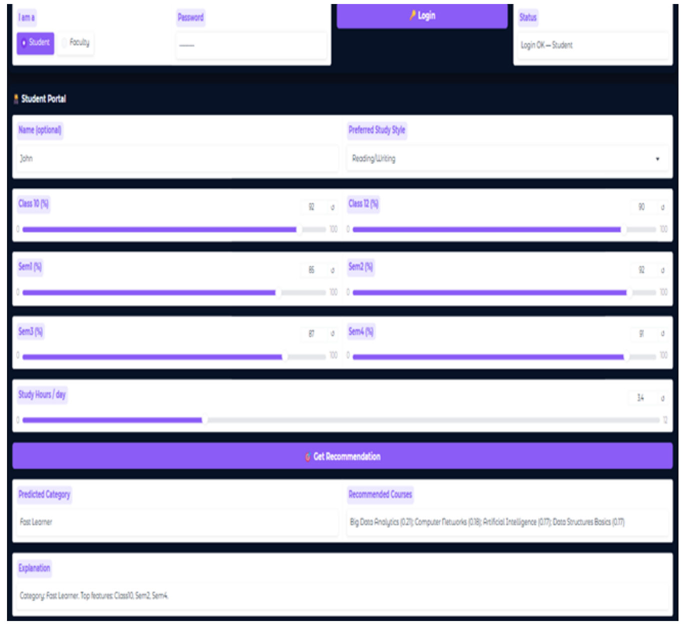

Fig. 7. Proposed interface for student Login

Faculties and mentors can use the tool to get the recommendation for all students, as shown in Figure 8 , faculty is allowed to login and provide the excel file containing students' information and download the excel file containing student classification and recommended courses information. They can also see bar chart showing count of each category of students and weights assigned by algorithm for each feature for each student.

VII. Conclusion and Future Work

This paper presents an Attention-Enhanced Heterogeneous Graph Neural Network that categorizes student learning profiles and recommends personalized courses. The framework provides a more realistic depiction of the context and relational complexity of learning settings by defining students, courses, and socio-academic characteristics as a type of node and attention between them as a relation-specific, compared to homogeneous or flat-feature models. The unified embedding space ensures that the network delivers classification and recommendation tasks in a single continuous process, and the attention mechanism provides the network with a clear understanding of the factors that underlie each prediction. Experimental results with real institutional data indicate that the method achieves 94.3 percent accuracy, exceeding the traditional machine-learning baseline and base GNN variants. The learner representations also exhibit evident divisions within groups of learners, ... which demonstrates the representational effectiveness of the model in encoding academic progressions and engagement behaviours.

The work can be improved in the future by developing it in several ways. To begin with, it would be better to test the model using multi-institutional and longitudinal student data to enhance its generalizability and encourage more of its application.

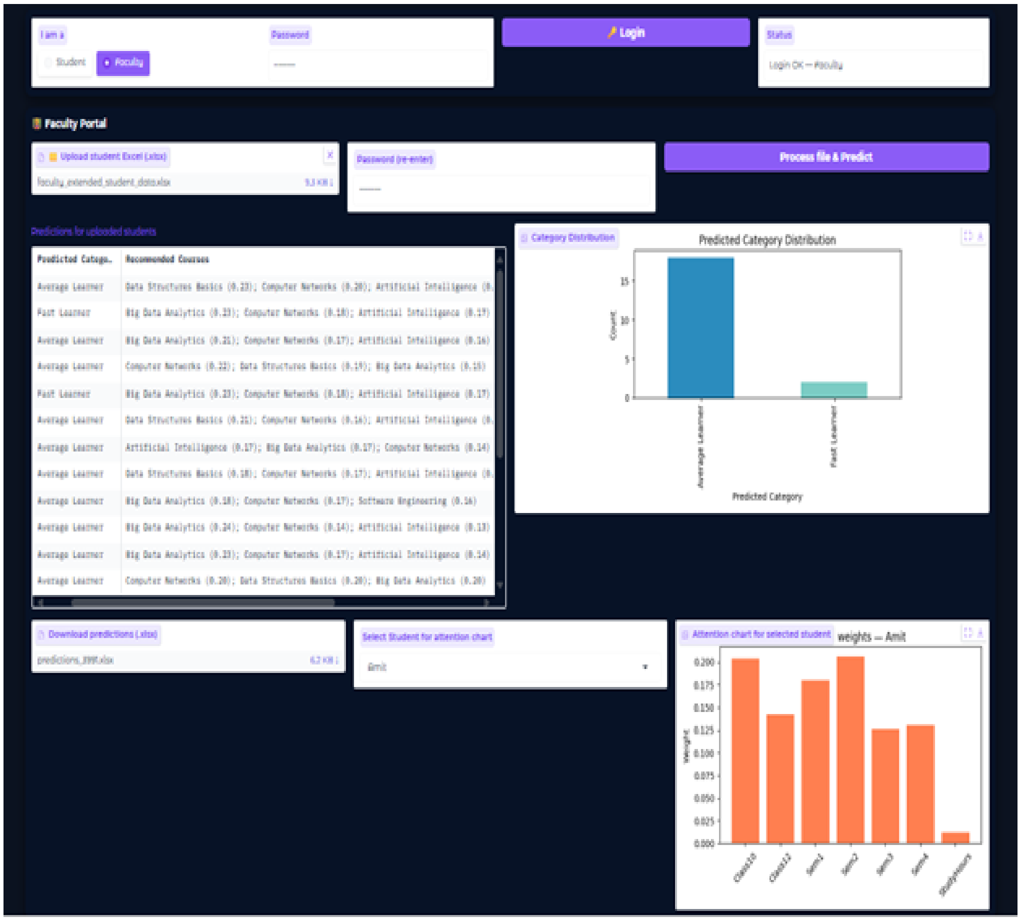

Fig. 8. Proposed interface for faculty login

TABLE VII. COMPARATIVE ANALYSIS OF PROPOSED METHOD

Study & Methodology	*Dataset Context*	*Reported Accuracy*	*Remarks*
A. Sharma and S. Kumar, “An improved student performance prediction model using ensemble learning [31]	University-level academic data (grades + demographics)	81%	Focused on ensemble classifiers; limited interpretability
M. Chen et al., “Student academic risk prediction using XGBoost and feature selection [32]	Large-scale MOOC dataset	78%	Strong performance on large data, but no relational modeling
D. Monteverde-Suárez et al., “Predicting students’ academic progress and related attributes in first-year medical students: an analysis with ANN and Naïve Bayes [33]	First-year medical school dataset	72–74%	Compared ANN and Naïve Bayes; accuracy moderate, no relational edges
S. Kumar and F. Janan, “Prediction of student’s performance using Random Forest Classifier [34]	Undergraduate student dataset	70%	Classical ML approach; lacks explainability
Proposed HGNN + Attention for Multimodal Student Profiling and Course Recommendation	Heterogeneous graph with academic, demographic, and contextual data	94.3%	Outperforms baselines, provides interpretability via attention weights, and extends to intelligent course recommendation

Introducing time- or sequence-sensitive learning models, including time-varying HGNNs, would be able to reflect dynamic fluctuations in the learner profiles across semesters. Also, the federated training would enable deployment across institutional boundaries with privacy. Lastly, the framework implementation in adaptive learning platforms may offer real-time academic feedback, enhancing individualized academic assistance on a large scale.

CONFLICT OF INTEREST

The authors declare that they have no known financial or personal relationships that could have appeared to influence the work reported in this paper. No external funding was received for this study.

REFERENCES

[1] A. Shahiri, W. Husain, and N. A. Rashid, "A review on predicting student's performance using data mining techniques," *Procedia Comput. Sci.*, vol. 72, pp. 414–422, 2015.

[2] R. Kumar, S. K. Singh, and R. K. Aggarwal, "Student performance prediction with optimized feature selection," *Procedia Comput. Sci.*, vol. 172, pp. 689–695, 2020.

[3] M. Li, X. Wang, Y. Wang, Y. Chen, and Y. Chen, "Study-GNN: A novel pipeline for student performance prediction based on multi-topology graph neural networks," *Sustainability*, vol. 14, no. 13, Art. 7965, Jun. 2022.

[4] Q. Hu and H. Rangwala, "Academic performance estimation with attention-based graph convolutional networks," *arXiv preprint arXiv:2001.00632*, Jan. 2020.

[5] Y. Wang, A. Ding, K. Guan, S. Wu, and Y. Du, "Graph-based ensemble machine learning for student performance prediction," *arXiv preprint arXiv:2112.07893*, Dec. 2021.

[6] L. Xiaochen, W. Wentao, Z. Zheng, and X. Bai, "Predicting student performance with graph structure density-based graph neural networks," in *Proc. Int. Conf. Mach. Learn. Intell. Comput. (ICMLIC)*, PMLR, vol. 245, pp. 25–32, 2024.

[7] X. Zhang, Y. Zhang, A. L. Chen, M. Yu, and L. Zhang, "Optimizing multi-label student performance prediction with GNN-TINet: A contextual multidimensional deep learning framework," *PLoS One*, vol. 20, no. 1, e0314823, Jan. 2025.

[8] T. H. Nguyen, T. B. Nguyen, and H. V. Nguyen, "Multi-view hypergraph neural networks for student academic performance prediction," *Expert Syst. Appl.*, vol. 215, 119514, 2023.

[9] S. X. Zhang, J. Bai, and Q. Wang, "Improving academic performance predictions with dual graph neural networks," *Appl. Intell.*, vol. 54, pp. 15879–15894, 2024.

[10] Y. Ektefaie, G. Dasoulas, A. Noori, M. Farhat, and M. Zitnik, "Multimodal learning with graphs," *Nat. Mach. Intell.*, vol. 5, no. 4, pp. 345–357, Apr. 2023.

[11] J. Yuan and K. Wang, "A survey on graph-based educational analytics," *Comput. Educ.*, vol. 207, 104401, 2023.

[12] F. Qian, L. Bai, L. Cui, M. Li, and H. Du, "Exploring the over-smoothing problem of graph neural networks for graph classification: An entropy-based viewpoint," in *Proc. AAAI Conf. Artif. Intell.*, 2025, pp. 19995–20003.

[13] T. Cai, Y. Jiang, M. Li, and C. Huang, "ML-GOOD: Multi-label graph out-of-distribution detection," in *Proc. AAAI Conf. Artif. Intell.*, 2025, pp. 15650–15658.

[14] P. Hooda, P. Mittal, B. D. Veerasamy, R. Bhatt, C. S. Lakshmi, S. Kamal, and P. K. Shukla, "Enhancing churn prediction through advanced machine learning techniques for modern education in computer science," Int. J. Mod. Educ. Comput. Sci. (IJMECS), vol. 17, no. 1, pp. 91–101, 2025, doi: 10.5815/ijmecs.2025.01.07.

[15] L. Bai, L. Cui, Z. Zhao, and M. Li, "AEGK: Aligned entropic graph kernels through continuous-time quantum walks," *IEEE Trans. Knowl. Data Eng.*, vol. 37, no. 3, pp. 1064–1078, 2025.

[16] S. Luan, C. Hua, Q. Lu, and L. Wu, "Revisiting heterophily for graph neural networks," in *Adv. Neural Inf. Process. Syst. (NeurIPS)*, 2023.

[17] N. Yanes, A. M. Mostafa, M. Ezz, and S. N. Almuayqil, "A machine learning-based recommender system for improving students learning experiences," *IEEE Access*, vol. 8, pp. 185577–185588, 2020, doi: 10.1109/ACCESS.2020.3036336.

[18] P. Kinnunen, R. McCartney, L. Murphy, and L. Thomas, "Through the eyes of instructors: A phenomenographic investigation of student success," in *Proc. 3rd Int. Comput. Educ. Res. Workshop (ICER'07)*, pp. 61–72, 2007, doi: 10.1145/1288580.1288589.

[19] V. Tam, E. Y. Lam, S. T. Fung, and W. W. T. Fok, "Enhancing educational data mining techniques on online educational resources with a semi-supervised learning approach," in *Proc. Int. Conf. Comput. Educ. (ICCE)*, Dec. 2015, pp. 203–206.

[20] N. A. Butt, Z. Mahmood, K. Shakeel, S. Alfarhood, M. Safran, and I. Ashraf, "Performance prediction of students in higher education using multi-model ensemble approach," *IEEE Access*, vol. 11, pp. 136091–136108, 2023, doi: 10.1109/ACCESS.2023.3336987.

[21] S. A. Hashmi, Y. Singh, and H. Bhardwaj, "Improving student classification model using Boruta feature selection technique," in *Communication and Intelligent Systems*, H. Sharma, V. Shrivastava, A. K. Tripathi, and L. Wang, Eds., Lecture Notes in Networks and Systems, vol. 1371. Singapore: Springer, 2025, pp. 165–176, doi: 10.1007/978-981-96-5723-0_13.

[22] Q. Al-Tashi, S. J. Abdul Kadir, H. M. Rais, S. Mirjalili, and H. Alhussian, "Binary optimization using hybrid grey wolf optimization for feature selection," *IEEE Access*, vol. 7, pp. 39496–39508, 2019, doi: 10.1109/ACCESS.2019.2906757.

[23] V. Hegde and H. S. Rao, "A framework to analyze performance of students in programming language using educational data mining," in *Proc. IEEE Int. Conf. Comput. Intell. Comput. Res. (ICCIC)*, 2017, pp. 1–4, doi: 10.1109/ICCIC.2017.8524244.

[24] P. Rojanavasu, "Educational data analytics using association rule mining and classification," in Proc. Joint Int. Conf. Digit. Arts, Media Technol. with ECTI Northern Sect. Conf. Electr., Electron., Comput. Telecommun. Eng. (ECTI DAMT-NCON), 2019, pp. 142–145.

[25] P. M. Arsad, N. Buniyamin, and J. L. A. Manan, "A neural network students' performance prediction model (NNSPPM)," in *Proc. IEEE Int. Conf. Smart Instrum., Meas. Appl. (ICSIMA)*, 2013, pp. 26–27, doi: 10.1109/ICSIMA.2013.6717966.

[26] Ş. Aydoğdu, "A new student modeling technique with convolutional neural networks: LearnerPrints," *J. Educ. Comput. Res.*, vol. 59, no. 4, pp. 603–619, 2021, doi: 10.1177/0735633120969216.

[27] K. Okoye *et al.*, "A text mining and statistical approach for assessment of pedagogical impact of students' evaluation of teaching and learning outcome in education," *IEEE Access*, vol. 11, pp. 15348–15360, Feb. 2023, doi: 10.1109/ACCESS.2023.3239779.

[28] V. U. Kumar, A. Krishna, P. Neelakanteswara, and C. Z. Basha, "Advanced prediction of performance of a student in a university using machine learning techniques," in *Proc. Int. Conf. Electron. Sustain. Commun. Syst. (ICESC)*, 2020, pp. 121–126, doi: 10.1109/ICESC48915.2020.9155557.

[29] K. Al Mayahi and D. M. Al-Bahri, "Machine learning based predicting student academic success," in *Proc. Int. Congr. Ultra Modern Telecommun. Control Syst. Workshops (ICUMT)*, Oct. 2020, pp. 264–268, doi: 10.1109/ICUMT51630.2020.9222435.

[30] E. G. Rincon-Flores, E. Lopez-Camacho, J. Mena, and O. O. Lopez, "Predicting academic performance with artificial intelligence (AI), a new tool for teachers and students," in *Proc. IEEE Global Eng. Educ. Conf. (EDUCON)*, Apr. 2020, pp. 1049–1054, doi: 10.1109/EDUCON45650.2020.9125141.

[31] A. Sharma and S. Kumar, "An improved student performance prediction model using ensemble learning," *IEEE Access*, vol. 10, pp. 105–115, 2022.

[32] M. Chen, Y. Zhang, L. Wang, and K. Huang, "Student academic risk prediction using XGBoost and feature selection," *Applied Sciences*, vol. 11, no. 23, pp. 11234–11247, Dec. 2021.

[33] D. Monteverde-Suárez, R. Garrote, and M. Ferrando, "Predicting students' academic progress and related attributes in first-year medical students: an analysis with ANN and Naïve Bayes," *BMC Medical Education*, vol. 24, no. 74, pp. 1–14, 2024.

[34] S. Kumar and F. Janan, "Prediction of student's performance using Random Forest classifier," *International Journal of Computer Applications*, vol. 183, no. 15, pp. 25–30, 2021.