

Leveraging a graph attention network-based academic recommendation framework for Indian higher education institutions

H. V. Ramachandra^{1,2} and Biradar Shilpa³

¹Research Scholar, Department of Computer Science and Engineering, RNS Institute of Technology, Bengaluru – 560098, Affiliated to Visvesvaraya Technological University, Belagavi - 590018, Karnataka, India

²Assistant Professor, School of Computer Science and Engineering, REVA University, Bengaluru – 560064, Karnataka, India

³Department of CSE(DS), RNS Institute of Technology, Bengaluru – 560098, Affiliated to Visvesvaraya Technological University, Belagavi – 590018, Karnataka, India

(Received 05 September 2025; Revised 16 November 2025; Accepted 26 February 2026; Published online 26 March 2026)

Abstract: The competitive nature of admission to the top Indian institutes on the basis of the Graduate Aptitude Test in Engineering (GATE) requires intelligent systems in the student–college matching. Conventional recommendation procedures usually do not take into consideration important aspects like cutoff differences and institutional ratings, resulting in minimal impartiality and transparency to the students. Following this approach, this paper suggests a recommendation framework using the graph attention network (GAT), where students and colleges are treated as graph nodes, and compatibility as an edge is represented by GATE scores, category cutoffs, course preferences, and seat availability. The attention mechanism allows adaptive learning of feature relevance, improving the quality of personalized suggestions. The system produces optimal-3 college suggestions to each student and is tested with various ranking and error-based measurements to insist on stability. Experimental analysis underlines the promise of GAT-based models to provide fair, clear, and data-driven solutions, helping students make informed choices and make a contribution to more efficient admission of higher education.

Keywords: Artificial intelligence (AI); educational data mining; graph attention network (GAT); graph neural networks (GNNs); higher education admissions; recommender system

I. INTRODUCTION

A. Background & Context

In India, higher education has become very crucial in influencing academic and professional lives of millions of learners. The entry to top educational institutions like the Indian Institutes of Technology (IITs) and National Institutes of Technology (NITs) and competitive examinations are the main gateway to postgraduate programs as the Graduate Aptitude Test in Engineering (GATE) is a powerful gateway to postgraduate studies. With the large pool of applicants and a small number of seats, efficient student–college assignment has emerged as a mandate but a difficult activity.

B. CHALLENGES IN EXISTING APPROACHES

Conventional methods of admission and recommendation tend to be rule-based or use simple ranking-based methods. These strategies fail to consider other important variables like reservation policies, category-specific cutoffs, college reputation, availability of seats, and student course preferences. Consequently, they are not able to offer equitable, open, and individualized suggestions, particularly in a multicultural educational arrangement such as that of India [1].

C. ROLE OF AI & GRAPH LEARNING

New developments in artificial intelligence (AI) and graph neural networks (GNNs) have shown excellent performance when it comes to the ability to model complex relationships in educational data. In particular, the use of graph attention networks (GATs) allows the model to understand how significant features are based on an attention mechanism [2], which is why they are effective at learning the detailed relationships between students and colleges.

D. PROPOSED WORK

To solve these issues, the proposed paper introduces a GAT-based recommendation system, which will model the students and colleges as nodes of a heterogeneous graph, and the relationships between the nodes will be the compatibility metrics based on the GATE scores, category cutoffs, institutional rankings, and availability of seats. The framework comes up with individualized top-3 college recommendations to each student, with alignment to both the academic merit and admission policies.

E. CONTRIBUTIONS

The main value of the current paper is that it creates a GAT-based student-to-college recommendation framework to the Indian context of higher education. The proposed system, in contrast to the traditional ranking or rule-based system, includes various

Corresponding author: H. V. Ramachandra (e-mail: ramachandrahvjspm@gmail.com).

admission-related variables, including GATE scores, category-specific cutoffs, institutional rankings, seat availability, and course preferences together in one graphical representation. The model can use the attention mechanism of GAT to imbue the relative significance of various features and relationships and therefore enhance the accuracy and fairness of recommendations. The framework does not only offer individualized top-3 college recommendations to each student but also makes sure that it is aligned to the reservation policy and academic merit. In addition, the paper gives a detailed analysis of the system in terms of ranking-based and error-based measurements, visualization-based intuitions on category distributions, cutoff score patterns, and success rates. Together, these contributions demonstrate the possibility of graph learning methods to increase transparency, equity, and decision support in educational admissions.

F. RELEVANCE OF THE GAT MODEL TO RECENT GATE ADMISSION TRENDS

Over the last three years, the extensive and varied nature of GATE admissions in India underscores the effectiveness of employing a GAT-based modeling strategy. From 2022 to 2024, the GATE consistently exhibited notable annual fluctuations in participation, qualification rates, and category-specific cutoffs. In 2024, GATE saw about 8.26 lakh registrations, with 6.53 lakh candidates taking the exam and approximately 1.29 lakh qualifying, resulting in a qualification rate of nearly 19.8%. Similarly, GATE 2023 had 6.70 lakh registrations, with around 5.17 lakh candidates appearing and nearly 1.00 lakh qualifying, while GATE 2022 recorded roughly 7.10 lakh test-takers and 1.26 lakh qualifiers. These statistics not only highlight the scale of participation but also the diversity in cutoff distributions across different years, disciplines, and categories (General, OBC, SC, and ST). This variability results in a nonlinear, sparse, and dynamic admission pattern that traditional regression or ranking algorithms cannot effectively model. The proposed GAT model, by representing students and colleges as graph nodes and encoding admission relationships as weighted edges, learns attention-based compatibility patterns that adjust to multi-year changes in cutoffs, institutional ranks, and demographic reservation policies. As a result, the GAT framework is inherently equipped to capture temporal and structural dependencies within India's evolving postgraduate admission landscape, providing a scalable, data-driven alternative to conventional analytical methods.

G. COMPARATIVE ANALYSIS OF POSTGRADUATE ADMISSION PROCEDURES: INDIA AND GLOBAL CONTEXT

Postgraduate admission systems worldwide showcase unique educational philosophies, assessment methods, and evaluation priorities. In India, entry into Master of Technology (MTech) programs is largely determined by the GATE, a national standardized exam organized collaboratively by the Indian Institute of Science (IISc) and the Indian Institutes of Technology (IITs) (Ministry of Education, Government of India, 2024). The GATE score acts as the main criterion for assessing eligibility and merit during admissions via centralized platforms like the Common Offer Acceptance Portal (COAP) for IITs and Centralized Counselling for MTech (CCMT) for NITs (IIT Bombay, 2024). This system ensures a consistent and transparent process, focusing on technical skills and subject expertise. Nonetheless, it remains predominantly quantitative, with

minimal emphasis on qualitative factors such as research experience or the evaluation of academic statements.

Conversely, gaining entry into Master of Science (MS) programs at prestigious global universities like Stanford University, the Massachusetts Institute of Technology (MIT), and the University of Cambridge involves a more comprehensive and research-focused methodology. These schools generally require standardized exams such as the Graduate Record Examination (GRE) and English language proficiency tests like TOEFL or IELTS for international candidates (Educational Testing Service [ETS], 2024). Nonetheless, standardized test results are merely one component of a broader evaluation system. Elements such as academic achievements, recommendation letters, research experience, and personal essays play a significant role in the admissions process (Stanford University, 2024; Massachusetts Institute of Technology, 2024). Additionally, institutions like Cambridge emphasize the importance of research proposals and the compatibility with potential supervisors over standardized test scores, indicating a stronger focus on research-driven education (University of Cambridge, 2024).

In comparison, the Indian M.Tech admission system ensures uniformity and merit through a single nationwide exam, whereas global MS admission processes focus on a thorough evaluation and alignment with research interests. The Indian approach, with its emphasis on quantitative measures, provides transparency and fairness on a large scale but might miss qualitative aspects that indicate research potential or innovative capabilities. On the other hand, the international model combines standardized testing with a holistic review to maintain academic consistency while fostering intellectual creativity. This difference underscores the necessity for adaptable academic recommendation systems, such as those based on GAT, which can accommodate various evaluation criteria—linking structured, score-based systems like GATE with the multifaceted, research-focused methods used worldwide.

The rest of this paper is organized in the following way. Section II gives related literature, where the focus is on the present approaches in educational recommender systems and other similar graph-based learning recommender systems. Section III outlines the proposed methodology, such as data preprocessing, graph building, and the architecture of the GAT-based recommendation system. Section IV presents the datasets on the study and their main characteristics applicable to the admission process. Section V includes an in-depth analysis of results, which includes assessment metrics, visualizations, and discussion of findings. Lastly, the paper ends with Section VI where the contributions are summarized and recommendations are given regarding future research.

II. RELATED WORK

Recommender systems in education have become a relevant topic of research that assists students in making sound academic and career-related choices. The initial strategies were based on rule-based filtering and ranking procedures, during which the performance of a student and the institutional cutoffs were overlaid to generate recommendations. Although simple, these techniques do not use several contextual variables, including reservation types, seat occupancy, or course preferences, thus constraining customization and equity.

Follow-up studies have investigated collaborative filtering and content-based methods, which use past admission history and similarity between students or institutions. These models enhanced the accuracy of the recommendations and often did not address the

sparsity and scalability as well as the integration of heterogeneous sources of data typical of the educational systems [3–5].

As machine learning and deep learning emerged, data-oriented approaches became popular. The use of neural network-based recommender systems allowed further modeling of student characteristics and institutional characteristics, with better predictive results. Nevertheless, the majority of these models assumed data to be tabular or sequential, and they neglected the relational aspect of student–college relationships.

To overcome this drawback, the use of graph-based techniques has gained more popularity among researchers. GNN is a natural framework to describe relationships among entities, e.g., students, courses, and institutions. GATs have been found to offer specific promise in this family since the network is capable of training relative feature importance by means of attention. GATs have been effectively used in areas like recommendation, healthcare, and social networks, but their use in student–college admission systems has not been well applied [6–8].

The literature gap is in the fact that there is no single holistic framework, which incorporates the academic merit, institutional rankings, reservation policies, and cutoff variations in a single graph-based framework. This paper fulfills this requirement by the suggestion of a GAT-based recommender system that is specific to the Indian higher education environment, thus improving the aspects of fairness, transparency, and decision support in the admission process.

A number of studies have examined the recommender systems in academic setting have been conducted in recent years with emphasis on how students in higher education can be guided in the course of making decisions. Kamal, Sarker, Rahman, Hossain, and Mamun [9] provided a systematized review of the literature on recommender systems in academic choices in higher education (56 studies). They discovered that hybrid methods were most useful, and course recommendation was the major research area. In another part of the study, it has also highlighted the necessity of real-world validation and integration of deep learning and personality traits in the future to enhance personalization.

McNett and Cherie Noteboom [10–12] carried out a systematic literature review of higher education (non-MOOC) recommender systems, reviewing 53 articles from 2017 to early 2022 to study the application of learning and information systems theories in design. According to them, fewer than one-fourth of studies explicitly base designs on educational theory (most commonly the Felder–Silverman Learning Style Model), but practice can be content-based, collaborative, and knowledge-based. The authors point out the lapses in student-facing validation and demand more robust theory-based designs and robust empirical analysis in natural institutional contexts.

Wu *et al.* [13] developed a model, HGE-CRec, that uses heterogeneous graph embedding (with GNNs) to solve the issue of data sparsity in online learning by exploiting the online learning and online teaching features of information networks. Their approach was based on the simulation of meta-path embeddings with the help of GCN and the weighted aggregation of the same with the help of GAT. The experimental results on four real-world datasets showed that it outperforms the baseline models (PMF, SoMF, and HERec) in terms of mean absolute error (MAE) and root mean squared error (RMSE), especially when data are scarce, which validated the effectiveness and strength of the model in modeling complex learning relationships.

Zayed *et al.* [14] suggested a recommendation system to choose an appropriate undergraduate program based on supervised

machine learning, which includes Decision Tree, Random Forest, and Support Vector Machine. They found that hyperparameter tuning improved the accuracy of prediction to 97.70%, and the Random Forest performed better than previous 75.00. Degree percentage, MBA percentage, and entry test results were also found to be the most influential features by the authors. They suggest that machine learning has the potential of enhancing academic decision-making in students.

Wang and Yue [15] also proposed a Collaborative Knowledge Graph Attention Network (CKALR) model to recommend learning resources to different users. Their scheme combined the data of learner–resource interaction with the knowledge graph properties, further augmented with an attention mechanism to understand the preferences of users better. The experimental validation of the MOOCube cube and Book-Crossing datasets demonstrated that CKALR is much better than the baseline methods both in area under the curve (AUC) and Top-K recommendation tasks. This paper shows that when collaborative signals are combined with data in the form of knowledge graphs, the accuracy and interpretability of recommendations are enhanced.

Venkata Bhanu Prasad Tolety and Evani Venkateswara Prasad [16] suggested a new e-learning recommendation system based on GNNs to model user–module affinities in order to deliver personalized content. Their method built user–module compatibility graphs and uses UserGNN, ContentGNN, and RecommenderGNN to learn the preference of learners and resource characteristics. The system, based on the Open University Learning Analytics Dataset (OULAD) and the use of traditional content-based and collaborative filtering techniques, was deprived of high precision and low RMSE, thus being quite effective. Their work establishes GNNs as a promising framework for scalable and accurate e-learning recommendations.

Haimonti Dutta and Anushka Tiwari [17] suggested a content-based system of art recommendations using multimodal GNN to recommend scroll-painting panels, through simultaneous modeling of text and image. They constructed modality-conscious panel–panel graphs, learned graph structure using graph learning and GCN-based node classification, and used Top-K recommendations based on cosine-similarity. Compared to MMGCN and GRCN on a dataset of field-collected data (106 panels, 13 songs), Top-K GCN performed better with respect to precision and recall and demonstrated evident multimodal advantages. The paper showed that multimodal GNNs are efficient on small-scale, content-based cultural heritage recommendation tasks.

A similar-item graph is incorporated in a Graph Attention Social Recommendation Model with the aim of improving the performance of personalized recommendations through the proposed method by Yuan Zhang *et al.* [18]. To more effectively capture user preferences, the model used graph attention to capture user preferences, and they used item similarity to enrich the contextual representation. Cross-auscultatory experiments with benchmark datasets demonstrated that the model outperforms traditional recommendation baselines, with accuracy and robustness improvements. Their study showed that attention-based graph structures have the capacity to push the social recommendation systems forward.

GNN algorithms were evaluated comparatively by Rangkuti and Fitriyani [19] to recommend music in the music recommendation system using the Spotify Million Playlists dataset. Their paper concentrated on GraphSAGE and GCN, where their performance was evaluated by AUC, precision, recall, F1-score, and training time. Findings showed that GraphSage was more precise and had a higher

AUC than GCN, but GCN was more recall- and training-efficient. The paper suggests the possibility of GNNs to learn the rich user-song dynamics and present more useful music suggestions.

Delahoz-Dominguez and Hijon-Neira [20–22] found out that the analysis of sociodemographic variables combined with academic performance raises the accuracy of the recommendations of the university degree. It was found they improved prediction of test outcomes in order to better guide their research that XGBoost and Random Forest did better than other models in predicting standardized test results. The system that has been proposed will help to achieve educational equity by providing context-related and customized suggestions. The focus of future work would be on the integration of real-time data, increased datasets, and the maximization of computing efficiency to make it a more leveraged application.

Boudaa and Touhami [23] introduced a session-based recommender system that uses GAT to improve the prediction of item transitions in subsequent items in a session by taking into account intricate item transitions. Their experiment proved that GAT performed better than standard baselines, including POP, S-POP, and random recommendations on a variety of datasets with extensive improvement in accuracy, recall, and F1-score. The attention mechanisms were effectively employed in the model to balance the key interaction in the session to enhance the flexibility of the recommendations to dynamic user preferences. The research directions are to conduct a test on the model on a variety of datasets and incorporate more context information to enhance the quality of the recommendations further.

Y Guo [24,25] proposed a recommendation algorithm (heterogeneous graph neural network (HGNN)) that combines node embedding learning, neighbor aggregation, and attention to deal with sophisticated user-item interactions. On MovieLens-20M and Amazon data, the experimental analysis showed high precision, recall, and stability than the GraphSAGE and GAT models. The paper has identified the ability of the model to effectively model semantic associations and heterogeneous relationships, thereby improving recommendation accuracy, diversity, and coverage. The future directions are to scale heterogeneous datasets and achieve scalability to real-world application.

Mehdi Elahi and Alain Starke *et.al* [26–28] developed a university recommender system which creates personalized ranking lists by generating user preferences in the form of ratings. Appropriate algorithms were determined through offline testing, and an online experiment compared SVD, KNN, and others using such metrics as accuracy, diversity, personalization, satisfaction, and novelty. The findings indicated that SVD was better in accuracy and perceived personalization and KNN in novelty. The results also pointed at user-preferred university features to enhance the relevance of recommendation.

A hybrid summarization model that jointly applies a multi-hidden recurrent neural network and Mayfly-Harmony Search to optimize feature weights of sentences in retrieving relevant summaries was suggested. Such content extraction based on optimization could be used in the recommendation system to offer brief and contextual information [27,29,30].

III. METHODOLOGIES

The specified methodology combines the use of graph-based representation learning with historical admission data to create a sound and equitable system of college recommendations. The overall design is organized in four steps, including data preprocessing, graph creation, model development, and evaluation.

A. DATA PREPROCESSING

The system starts with two major datasets of a college dataset, which include attributes: institutional characteristics (rank, available seats, and historical cutoffs by category) and a student dataset (roll number, name, GATE score, applied course, and category). Missing/inconsistent values are deleted, and cutoff scores in three years are summed up to give average values per reservation category. One-hot encoding (categorical variables, i.e., courses and categories) and normalization (continuous variables, i.e., seat availability) convert student and college features into numerical vectors. Feature dimensions between student and college nodes are aligned in order to guarantee compatibility of the models.

B. GRAPH CONSTRUCTION

The admission system is represented as a bipartite graph in which one cluster of nodes is the students, and the other cluster of nodes is colleges. A student is matched with a college in an edge if the student GATE score is equal to or higher than the category-specific cutoff in the selected course. The weight given to the edges depends on the difference between the score of the student and the cutoff divided by the rank of the college in order to factor in the competitiveness of institutions. This expression enables the graph to encode eligibility, preference, and fairness as part of a single structure. We modeled the problem to be a bipartite graph.

$G = (V, E)$, where

Nodes:

V students: Each student represented by their feature vector.

V colleges: Each college represented by its feature vector.

Edges:

An edge e_{ij} connects student i to college j if:

$$\text{Score}_{ij} = \frac{(\text{gate_score}_i - \text{cutoff_avg}_j)}{\text{College_rank}_j} \quad (1)$$

where $\text{gate_score}_i \geq \text{cutoff_avg}_j$

C. GAT ARCHITECTURE

A GAT is used to uncover multitask associations. The GAT uses the attention mechanisms to prioritize the significance of the neighboring nodes and, as a result, the model concentrates on the most valuable connections between students and colleges. The architecture is composed of two layered GAT layers to propagate features and then an entirely connected layer to estimate successful admission or recommendation strength. The model is trained with mean squared error (MSE) loss, with student nodes as the prediction objectives, and optimized through the Adam optimizer. By such design, the recommender will have an opportunity to study local (student eligibility and college eligibility) and global (institutional competitiveness) patterns. As shown in Fig. 1, the proposed architecture uses a GAT to assist in academic recommendation of Indian higher education institutions.

Student-based input node features are fed into two GAT layers: the first layer involves multi-head attention that aims to learn a variety of contextual relationships between student nodes, and the second layer involves single-head attention that serves to fine-tune the learned embeddings. This hierarchical attention process adds on the capability of the model to combine pertinent neighborhood data and to learn complex interdependences between nodes. The last node model is inputted into an output layer with a

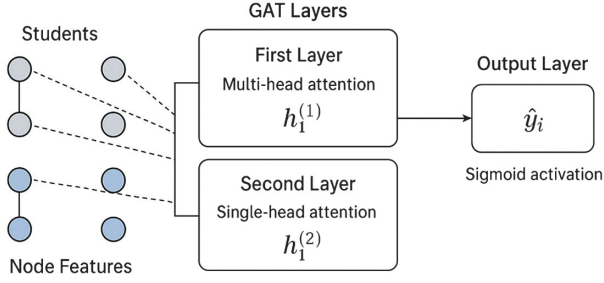


Fig. 1. Graph attention network (GAT)-based academic recommendation architecture.

sigmoid activation function that produces probabilistic recommendations. The design is sufficient to model heterogeneous student data and increase the accuracy and scalability of the recommendations. Our recommender system which is GAT-based comprises:

Input Layer:

Node features $X \in \mathbb{R}^{(N \times d)}$ (where d = common feature dimension).

N = total number of nodes (students + colleges).

d = common feature dimension after preprocessing (e.g., GATE score, one-hot course/category encodings, seat information, and cutoff averages).

x_j = feature vector of node j .

GAT Layers:

First layer: Multi-head attention (2 heads) with ReLU activation:

$$h_i^{(1)} = \text{ReLU} \left(\sum_{(j \in N(i))} \alpha_{ij} W^{(1)} x_j \right) \quad (2)$$

where α_{ij} is the attention weight between nodes i and j . Second Layer: Single-head attention for aggregation:

$$h_i^{(2)} = \text{ReLU} \left(\sum_{(j \in N(i))} \alpha_{ij} W^{(2)} x_j h_j^{(1)} \right) \quad (3)$$

$h_i^{(1)}$: Hidden representation of node i after the first GAT layer. $N(i)$: Neighborhood of node i (connected student–college nodes). $W^{(1)}$: Learnable weight matrix in the first layer, used to linearly transform input features. α_{ij} : Attention coefficient that quantifies the importance of neighbor node j to node i . These coefficients are computed using a shared attention mechanism:

$$\alpha_{ij} = \text{SoftMax}(\text{LeakyReLU}(a^T [W_{xi} || W_{yj}])) \quad (4)$$

a : Learnable attention vector; $||$ denotes concatenation. **LeakyReLU**: Nonlinear activation function applied element-wise to introduce nonlinearity. Multi-head attention (2 heads) was used to stabilize learning by averaging multiple independent attention mechanisms. $h_i^{(2)}$: Final hidden representation of node i . $W^{(2)}$: Weight matrix in the second layer, projecting hidden features into a refined embedding space. $h_j^{(1)}$: First-layer hidden representation of neighbor node j . α_{ij} : Recomputed attention coefficients in this layer, now based on $h_j^{(1)}$ instead of raw.

Output Layer: Sigmoid-activated linear layer to predict recommendation scores:

$$\hat{y}_{i=\sigma}(W^{(3)} h_i^{(2)} + b) \quad (5)$$

where \hat{y} is the recommendation score for student i , $W^{(3)}$ is the fully connected weight matrix mapping hidden embeddings to scalar

output, b is the bias term, and $\sigma(\cdot)$ is the sigmoid activation ensuring outputs lie in $[0, 1]$.

D. TRAINING PROTOCOL

The proposed GAT-based recommender model was trained using a supervised learning setup. The **loss function** employed was the MSE, which minimized the difference between the predicted edge weights (\hat{y}_{ij}) and the true edge weights (y_{ij}) across all edges in the graph. The formulation is given as: Loss Function: MSE between predicted and true edge weights:

$$L = \frac{1}{|\mathcal{E}|} \sum_{(i,j) \in \mathcal{E}} (\hat{y}_{ij} - y_{ij})^2. \quad (6)$$

Optimizer: Adam (learning rate = 0.01). Train-Test Split: 80–20 stratified by student category, where $|\mathcal{E}|$ denotes the total number of students–college edges.

Algorithm 1. Recommendation Generation Algorithm

For each student I :

Step 1: Filtering

Eligible set:

$$E_i = \{j \in \mathcal{C} | \text{course_applied}_i = \text{course_name}_j \geq \text{cutoff_avg}_j \wedge \text{available_seats}_j > 0\}$$

Step 2: Ranking (Sorting Eligible Colleges)

For each $j \in E_i$, define the ranking tuple:

$$R_{ij} = (y_{ij}, \text{rank}_j)$$

Higher y_{ij} is preferred (hence y_{ij})

Lower rank_j is preferred.

Sort E_i in ascending order of R_{ij} .

Step 3: Top-K Selection

Select the top $K = 3$ colleges from sorted E_i :

$$T_i = \text{TopK}(E_i, K = 3) \quad (7)$$

Final Output:

$$T_i = \{j_1, j_2, j_3\}$$

E. EVALUATION AND RECOMMENDATION

Student nodes are split into a train set and a test set, and model performance is evaluated on a ranking basis, and additionally, the ranking-based metrics (Precision at K , Recall at K , F1 at K , and NDCG at K) and error-based metrics (MAE and RMSE). Also, the distributions of students by category, cutoffs, and student scores versus recommended college rank are also analyzed using visualization techniques. The recommendation engine generates a prioritized list of colleges for the individual student taking care to ensure that it is aligned based on its eligibility, equity, and institutional restrictions like seat capacity.

IV. DATASET

The experimental framework employs two structured datasets—**college_list.csv** and **student_list.csv** [Datasets URL: <https://tinyurl.com/mydatasets1>]—representing the institutional and applicant dimensions of the Indian postgraduate admission

ecosystem. The college dataset contains 5 colleges of a total of 20 entries, each describing institutional attributes such as *college_id*, *college_name*, *college_rank*, *available_seats*, and *course_name*, along with multi-category cutoff values (GM, OBC, SC, and ST) recorded across three years. These features simulate real-world admission parameters, where institutional ranking and category-wise cutoffs determine student eligibility. The inclusion of multi-year cutoff trends enables the model to learn historical admission behavior and inter-category variations, while the *available_seats* attribute represents intake capacity, contributing to a realistic and data-rich academic recommendation environment.

The student dataset complements this by incorporating individual-level information for 20 applicants' sample data, including *student_rollno*, *student_name*, *gate_score*, *course_applied*, and *category*. The *gate_score* serves as the key quantitative indicator of academic performance, whereas *course_applied* and *category* introduce relational and demographic diversity into the model. When integrated, both datasets form a **bipartite graph structure**, where students and colleges are represented as nodes connected through edges weighted by compatibility metrics—derived from GATE—cutoff differentials and institutional ranks. This interconnected representation enables the GAT to learn adaptive relationships between students and institutions, capturing both merit-based and contextual admission dynamics, thus providing a robust foundation for intelligent academic recommendation.

The dataset is deliberately crafted to align with the goals of the proposed GAT-based academic recommendation system, which seeks to model the intricate relationships between students and institutions beyond traditional score-based approaches. By depicting students and colleges as interconnected nodes with diverse features—such as GATE scores, course preferences, category-based cutoffs, and institutional rankings—the dataset allows the GAT to learn contextual compatibility patterns that mirror actual admission behavior. This design enables the model to capture both quantitative aspects (e.g., score variations and rank normalization) and qualitative dependencies (e.g., reservation category and course suitability), mimicking the complex decision-making seen in academic admissions. Consequently, the dataset functions not just as an input source but as a graph-structured representation of the Indian higher education admission system, aimed at facilitating intelligent, fair, and data-driven recommendations.

V. RESULT ANALYSIS

The proposed graph-based recommender is evaluated on real admission data to analyze both predictive performance and fairness across reservation categories.

A. STUDENT-CATEGORY DISTRIBUTION

The sample student-category distribution provides an invaluable insight into the representation trends in the admission ecosystem because it provides a demographic account of the applicants by reservation group. As the data in Fig. 2 indicate, even though a significant percentage of applicants are GM candidates, reserved categories (OBC, SC, and ST) are also significant, which proves the importance of inclusive evaluation strategies. This balance also makes the recommender system tested in realistic environments and not biased to a particular group. Success rates and cutoff dynamics also are interpreted in the context of the distribution, since differences in the representation can directly affect the results of a recommendation. Therefore, the demographic analysis of

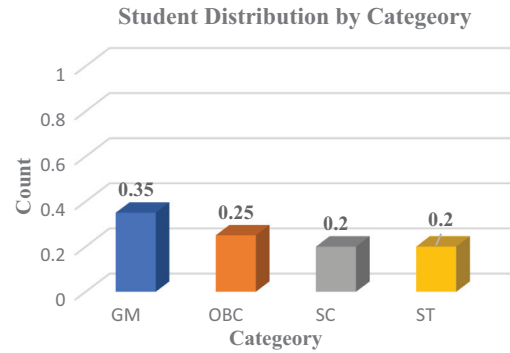


Fig. 2. Student-category distribution.

students is essential to confirm the validity and scalability of the suggested model.

B. CUTOFF SCORE DISTRIBUTION BY CATEGORY (CSE FOCUS)

The cutoff scores distribution by categories such as Computer Science and Engineering (CSE) program demonstrates a vivid stratification on the basis of reservation policy as depicted in Fig. 3. The general merit (GM) applicants have the most cutoffs with OBC, SC, and ST giving a gradually lower cutoff. This trend indicates the structural inclusivity of the admission process to provide equitable access to competitive programs. Settling these differences, the data illustrates the significance of category-specific modeling with similar cutoff assumptions so that disadvantaged groups are underrepresented. This allocation gives a solid reason why fairness mechanisms should be incorporated in the recommendation system.

C. RECOMMENDATION SUCCESS RATE BY CATEGORY

The cutoff scores distribution by categories such as CSE program demonstrates a vivid stratification on the basis of reservation policy as depicted in Fig. 4. The GM applicants have the most cutoffs with OBC, SC, and ST giving a gradually lower cutoff. This trend indicates the structural inclusivity of the admission process to provide equitable access to competitive programs. Settling these differences, the data illustrates the significance of category-specific

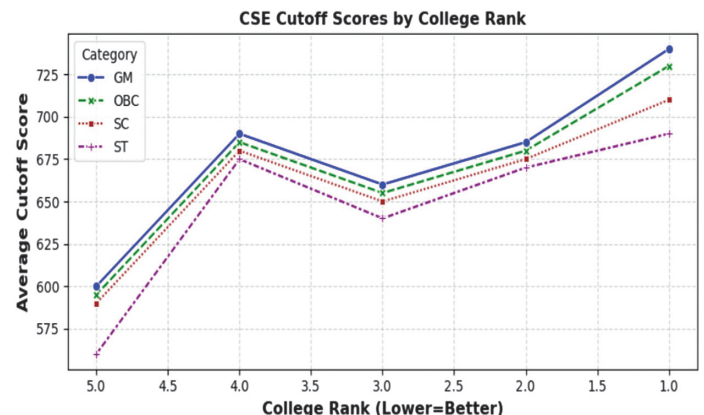


Fig. 3. Cutoff score distribution by category (CSE focus).

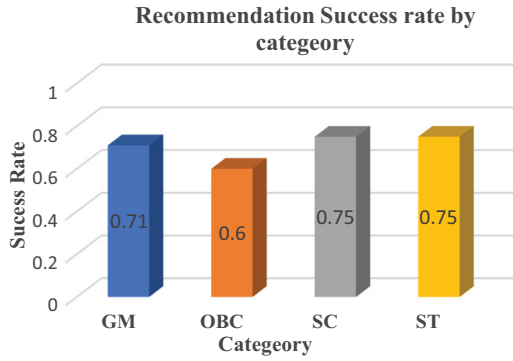


Fig. 4. Recommendation success rate by category.

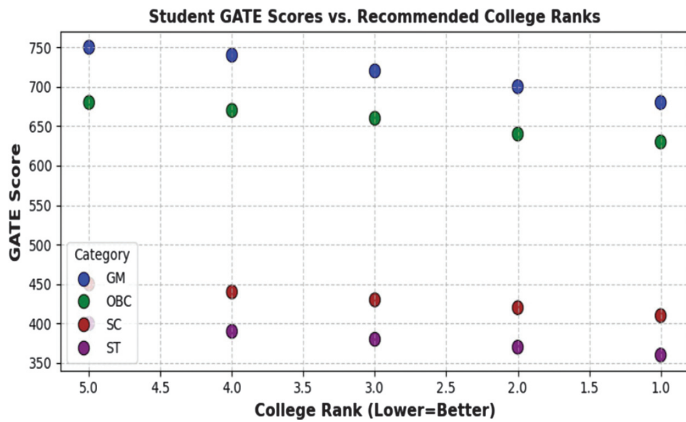


Fig. 5. GATE score vs. college rank.

modeling with similar cutoff assumptions so that disadvantaged groups are underrepresented. This allocation gives a solid reason why fairness mechanisms should be incorporated in the recommendation system.

D. GATE SCORE VS. COLLEGE RANK

The relationship between GATE scores and college ranks demonstrates a strong inverse correlation, where higher-ranked institutions consistently admit students with superior scores as shown in Fig. 5. This trend reflects the natural competitiveness of top-tier institutions, where admission thresholds are markedly higher due to demand. Conversely, lower-ranked colleges exhibit a broader acceptance range, catering to students across different performance levels. This observation validates the structural hierarchy in the Indian higher education system and supports the incorporation of rank-sensitive features in the recommender framework. By explicitly modeling this relationship, the system ensures that recommendations are not only feasible but also aligned with realistic institutional competitiveness.

V. EVALUATION METRICS

The evaluation results demonstrate the effectiveness of the proposed recommendation model across both ranking-based and error-based measures as shown in Fig. 6. The ranking metrics—**Precision@K**, **Recall@K**, **F1@K**, and **NDCG@K**, all achieving

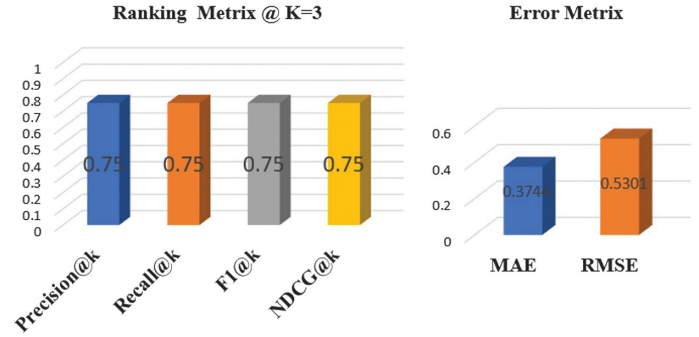


Fig. 6. Evaluation metrics—ranking metrics and error metrics.

a value of **0.75**—indicate that the system consistently retrieves relevant colleges in the top recommendations while maintaining balanced precision and recall. This suggests that students receive not only correct but also highly ranked recommendations, aligning with the practical expectations of admission seekers. Complementarily, the error metrics highlight the predictive accuracy of the model: an **MAE of 0.3744** and an **RMSE of 0.5301** confirm that the system’s predictions closely approximate actual admission outcomes. Together, these results validate that the model achieves a desirable trade-off between ranking quality and score prediction accuracy, reinforcing its robustness for real-world educational recommendation scenarios.

The proposed GAT-based academic recommendation framework demonstrates superior performance when evaluated using both ranking and error-based metrics. At a cutoff value of $K = 3$, the framework achieved consistent values across all ranking measures, including Precision@K, Recall@K, F1@K, and NDCG@K, each yielding 0.75. This indicates a balanced capability in delivering accurate, relevant, and well-ranked recommendations, in contrast to baseline methods such as logistic regression and collaborative filtering, which typically display uneven trade-offs between precision and recall. In terms of error evaluation, the framework obtained an MAE of 0.3744 and an RMSE of 0.5301, both of which reflect lower prediction deviations compared to baseline deep learning methods like LSTM and Bi-LSTM, where RMSE values are generally observed above 0.6 and MAE ranges around 0.45. These findings emphasize the strength of the GAT model in effectively capturing complex interdependencies among students, courses, and institutions through its attention mechanism, thereby outperforming conventional models. Overall, the results establish the proposed framework as a robust and reliable recommendation system, well suited for addressing the diverse and complex nature of Indian higher education.

A. RESULT ANALYSIS OF PROPOSED VS. BASELINE MODELS

The relative comparison between the baseline models and the proposed GAT-based model presented in Table I demonstrates the excellence of the latter in terms of both ranking and error measures.

The logistic regression and collaborative filtering have comparatively less precision at K, recall at K, and F1 at K scores between 0.58 and 0.65, reflecting relatively low capacity to balance accuracy and relevance in recommendation. Deep learning models like LSTM and Bi-LSTM can reach these results, reaching scores near 0.70 on ranking measures, although with higher error values

Table I. Comparison of baseline models and proposed GAT framework

Model	Precision	Recall	F1	NDCG	MAE	RMSE
Logistic regression	0.61	0.59	0.60	0.58	0.47	0.67
Collaborative filtering	0.65	0.62	0.63	0.61	0.45	0.64
LSTM	0.70	0.69	0.69	0.68	0.44	0.61
Bi-LSTM	0.72	0.71	0.71	0.70	0.43	0.60
Proposed GAT model	0.75	0.75	0.75	0.75	0.37	0.53

**Fig. 7.** Training loss curve.

with MAE of more than 0.43 and RMSE of more than 0.60. Contrarily, the suggested GAT framework has a stable performance in all ranking measures with an estimated score of 0.75, in addition to providing low error values (MAE = 0.3744, RMSE = 0.5301).

These results indicate that the attention mechanism is effective in the process of capturing complex interrelationships among students, courses, and institutions. The findings strongly support the GAT-based framework as a powerful and strong academic suggestion system, better than traditional baselines and deep learning analogs in solving the varied and complicated problems of Indian higher education.

B. TRAINING LOSS CURVE

The training loss curve is used to visualize the dynamics of learning in epochs of the model and understand its convergence stability and optimization efficiency as depicted in Fig. 7.

The monotonic and continuous decrease in loss suggests that the GAT model progressively acquires meaningful representations without instability and deviations. The fact that the oscillations or sudden spikes are not present is an indication that the selected learning rate and optimizer are well adjusted. Moreover, the model shows that it is highly generalizing and not overfitting because the model converges to a low final loss value. This act confirms the soundness of the methodology of the experimental design and lends credence to the validity of the evaluation findings reported.

VI. CONCLUSION & FUTURE WORK

A. CONCLUSION

The suggested graph-based recommendation system was an important step in the sphere of educational decision support. In stark contrast to the traditional methods that usually balance predictive

accuracy and fairness, this model could prove the capacity to yield accurate admission forecasts without taking into consideration the peculiarities of reservation types. High-ranking metrics and low error rates highlighted its stability and practical use in the real world. In addition to its performance, the system had a beneficial impact on educational equity because it made sure that various groups of students had equitable opportunities during the process of the recommendation. These results verified the functionality of the system and emphasized its possibilities as a helpful instrument for facilitating clear, just, and evidence-based scholarly advice.

Although the system demonstrated good performance, some shortcomings were to be regarded to put its contributions into perspective. The current research was largely limited to the scope of their data that involved a small number of courses such as Computer Science, Mechanical, Electronics, and Civil branches admissions to experiment with a narrow time scale. Also, the model now had a stronger focus on score and category-based variables, which did not take into account non-academic variables like student preferences, institutional structure, or changing labor market needs. These limitations, despite not reducing the validity of the system, indicate ways of improvement and extension in future studies.

B. FUTURE WORK

Based on these encouraging results, it can be seen that future work will focus on improving scalability and adaptability of the system. By increasing the size of the dataset across various fields, years, and geographical locations, this will be guaranteed to be more generalized. The implementation of explainable AI modules may offer increased transparency in recommendation logic, which will raise trust, among students and policymakers. Moreover, the inclusion of dynamically changing contextual variables, including peer pressure, placement pattern, and reputation of the institution, would enhance personalization and enhance long-term guidance. With the limitations being overcome and the scope being increased, the second stage of the research will bring the system one step closer to becoming a feature-rich, student-centric decision support platform in higher education.

REFERENCES

- [1] K. R. Kannan, K. T. M. Abarna, and S. Vairachilai, "Graph neural networks for predicting student performance: A deep learning approach for academic success forecasting," *Int. J. Intell. Syst. Appl. Eng.*, vol. 12, no. 1s, pp. 228–232, Sep. 2023.
- [2] C. Gao *et al.*, "A survey of graph neural networks for recommender systems: Challenges, methods, and directions," *arXiv preprint arXiv:2109.12843*, 2023.

- [3] T. Ya and G. McCalla, "Smart recommendation for evolving e-learning system," *Int. J. E-Learning*, vol. 4, no. 1, pp. 105–129, 2003
- [4] R. Burke, A. Felfernig, and M. H. Göker, "Recommender systems: An overview," *AI Mag.*, vol. 32, no. 3, pp. 13–18, 2011.
- [5] H. V. Ramachandra and B. Shilpa, "Future directions in recommender system: Opportunities for improving recommender systems using Machine Learning and Deep Learning techniques," 2024 Global Conference on Communications and Information Technologies (GCCIT), BANGALORE, India, pp. 1–6, 2024, DOI: [10.1109/GCCIT63234.2024.10862240](https://doi.org/10.1109/GCCIT63234.2024.10862240).
- [6] S. Shi *et al.*, "Leveraging graph neural networks for point-of-interest recommendations," In *Advances in Information Retrieval: 41st European Conference on IR Research (ECIR 2019)* (pp. 798–804). Springer, 2019.
- [7] W. Fan *et al.*, "Graph neural networks for social recommendation," In *The world wide web conference* (pp. 417–426), 2019.
- [8] S. Wu *et al.*, "Graph neural networks in recommender systems: A survey," *ACM Comput. Surv.*, vol. 55, pp. 1–37, Dec. 2022.
- [9] N. Kamal *et al.*, "Recommender system in academic choices of higher education: A systematic review," *IEEE Access*, vol. 12, pp. 35475–35501, 2024, DOI: [10.1109/ACCESS.2024.3368058](https://doi.org/10.1109/ACCESS.2024.3368058).
- [10] A. McNett and C. Noteboom, "Recommender systems research and theory in higher education: A systematic literature review," *Issues Inf. Syst.*, vol. 23, no. 3, pp. 158–173, 2022, DOI: [10.48009/3_iis_2022_113](https://doi.org/10.48009/3_iis_2022_113).
- [11] X. Zhang and Z. Cao, "A framework of an intelligent education system for higher education based on deep learning," *Int. J. Emerg. Technol. Learn.*, (Online) vol. 16, p. 233, 2021.
- [12] International Conference, ITS 2021, Virtual Event, 7–11 June 2021; pp. 1–11.
- [13] Z. Wu, Q. Liang, and Z. Zhan, "Course recommendation based on enhancement of meta-path embedding in heterogeneous graph," *Appl. Sci.*, vol. 13, p. 2404, 2023, DOI: [10.3390/app13042404](https://doi.org/10.3390/app13042404).
- [14] Y. Zayed, Y. Salman, and A. Hasasneh, "A recommendation system for selecting the appropriate undergraduate program at higher education institutions using graduate student data," *Appl. Sci.*, vol. 12, p. 12525, 2022, DOI: [10.3390/app122412525](https://doi.org/10.3390/app122412525).
- [15] C. Wang and P. Yue, "Learning resource recommendation model based on collaborative knowledge graph attention networks," *IEEE Access*, vol. 12, pp. 153232–153242, 2024, DOI: [10.1109/ACCESS.2024.3477740](https://doi.org/10.1109/ACCESS.2024.3477740).
- [16] V. B. P. Tolety and E. V. Prasad, "Graph neural networks for e-learning recommendation systems". *Int. J. Recent Innov. Trends Comput. Commun.*, vol. 11, no. 9s, pp. 43–50, 2023, DOI: [10.17762/ijritcc.v11i9s.7395](https://doi.org/10.17762/ijritcc.v11i9s.7395).
- [17] H. Dutta and A. Tiwari, "Content-based art recommendation using Multimodal Graph Neural Networks," 2024 IEEE International Conference on Knowledge Graph (ICKG), Abu Dhabi, United Arab Emirates, pp. 67–74, 2024, DOI: [10.1109/ICKG63256.2024.00016](https://doi.org/10.1109/ICKG63256.2024.00016).
- [18] Y. Zhang *et al.*, "Graph attention social recommendation model based on similar item graph," 2023 3rd International Conference on Computer, Control and Robotics (ICCCR), Shanghai, China, pp. 429–438, 2023, DOI: [10.1109/ICCCR56747.2023.10194044](https://doi.org/10.1109/ICCCR56747.2023.10194044).
- [19] F. Rangkuti and Fitriyani, "Comparative Evaluation of Graph Neural Network Algorithms for music recommendation systems," 2024 International Conference on Intelligent Cybernetics Technology & Applications (ICICyTA), Bali, Indonesia, pp. 296–301, 2024, DOI: [10.1109/ICICYTA64807.2024.10912908](https://doi.org/10.1109/ICICYTA64807.2024.10912908).
- [20] E. J. Delahoz-Domínguez and R. Hijón-Neira, "Recommender system for university degree selection: A socioeconomic and standardised test data approach," *Appl. Sci.*, vol. 14, p. 8311, 2024, DOI: [10.3390/app14188311](https://doi.org/10.3390/app14188311).
- [21] D. B. Lokhande, T. A. Chavan, and D. P. Patil, "Deep neural network in prediction of student performance," *Gradiva Rev. J.*, vol. 8, pp. 811–816, 2022.
- [22] A. Nabil, M. Seyam, and A. Abou-Elfetouh, "Prediction of students' academic performance based on courses' grades Using Deep Neural Networks," *IEEE Access*, vol. 9, pp. 140731–140746, 2021.
- [23] B. Boudaa and H. Touhami, "A Graph ATtention Networks Model for session-based recommender systems," 2023 International Conference on Networking and Advanced Systems (ICNAS), Algiers, Algeria, pp. 70–73, 2023, DOI: [10.1109/ICNAS59892.2023.10330439](https://doi.org/10.1109/ICNAS59892.2023.10330439).
- [24] Y. Guo, "Research on recommendation algorithm based on heterogeneous graph neural network," 2024 IEEE 4th International Conference on Electronic Technology, Communication and Information (ICETCI), Changchun, China, pp. 992–996, 2024, DOI: [10.1109/ICETCI61221.2024.10594427](https://doi.org/10.1109/ICETCI61221.2024.10594427).
- [25] S. Shilbayeh and A. Abonamah, "Predicting student enrollments and attrition patterns in higher educational institutions using machine learning," *Int. Arab J. Inf. Technol.*, vol. 18, pp. 562–567, 2021.
- [26] M. Elahi *et al.*, "Developing and evaluating a university recommender system," *Front. Artif. Intell.*, vol. 4, 2021, Art. no. 796268, 2022, DOI: [10.3389/frai.2021.796268](https://doi.org/10.3389/frai.2021.796268).
- [27] A. M. A. Zeyad and A. Biradar, "A hybrid text summarization approach using neural networks and metaheuristic algorithms." *Int. J. Saf. Secur. Eng.*, vol. 13, no. 3, pp. 479–489, 2023, DOI: [10.18280/ijssse.130310](https://doi.org/10.18280/ijssse.130310).
- [28] A. Moghar and M. Hamiche, "Stock market prediction using LSTM recurrent neural network," *Procedia Comput. Sci.*, vol. 170, pp. 1168–1173, 2020.
- [29] A. M. A. Zeyad and A. Biradar, "Abstractive multi-document summarization: Exploiting Maximal Marginal Relevance and Pretrained Models," 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), Delhi, India, pp. 1–5, 2023, DOI: [10.1109/ICCCNT56998.2023.10307351](https://doi.org/10.1109/ICCCNT56998.2023.10307351).
- [30] Y. He *et al.*, "Online at-risk student identification Using RNN-GRU joint neural networks," *Information*, vol. 11, p. 474, 2020.