International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(11s) 2026 pp. 1307-1314
https://doi.org/10.70917/ijcisim-2026-3868
Article

Knowledge Graph-Enhanced Course Recommendation System: A Multi-Relational Semantic Approach

Win Mathew John[1], T. Ramaprabha[2], Kochumol Abraham[3]*

[1]Department of Computer Science, Nehru Arts and Science College, Coimbatore, India.
[2]Department of Computer Science, Nehru Arts and Science College, Coimbatore, India.
[3]PG Department of Computer Applications, Marian College Kuttikanam Autonomous.
Email: **kochumol.abraham@mariancollege.org**
Corresponding Author: Kochumol Abraham

Abstract: —India's New Education Policy introduces unprecedented flexibility in higher education through multidisciplinary programs, multiple entry-exit points, and credit-based academic mobility. This paradigm shift necessitates intelligent systems to guide students through complex course selection decisions. This paper presents a novel Knowledge Graph (KG) enhanced course recommendation framework specifically designed for the New Education Policy compliance. We construct an Educational Knowledge Graph comprising 238 entities (122 courses, 110 skills, 6 career pathways) with multi-relational edges capturing course-skill associations, prerequisite dependencies, and career alignment mappings. Our recommendation algorithm leverages graph traversal techniques combined with semantic similarity measures to generate personalized course suggestions. Experimental evaluation on 28,785 real student trajectories from the Open University Learning Analytics Dataset (OULAD) demonstrates that our KG-enhanced approach achieves significant improvements: 30.2% over popularity-based baselines and 65.1% over random selection in NDCG@5. The study provides explainable recommendations, including skill-gap analysis and career-alignment justifications, supporting the vision of outcome-based, student-centric education. We release our implementation and processed datasets to facilitate reproducibility and further research in educational recommendation systems.
Keywords: —Knowledge Graph, Course Recommendation, Educational Data Mining, Personalized Learning, Semantic Web, Higher Education

1. INTRODUCTION

The New Education Policy represents the most significant reform in Indian higher education since independence, introducing flexible degree structures with multiple entry and exit points, credit accumulation and transfer mechanisms through the Academic Bank of Credits (ABC), and emphasis on multidisciplinary learning [1]. These reforms create unprecedented complexity in academic planning, where students must navigate extensive course catalogs while aligning their choices with career aspirations, skill development goals, and prerequisite requirements.

Traditional course registration systems offer limited guidance, typically presenting alphabetical course lists or department-wise catalogs without semantic understanding of course relationships [2]. This limitation particularly affects first-generation learners, who account for approximately 50% of enrollments in Indian higher education, according to the All-India Survey on Higher Education (AISHE) 2021 [3]. These students often lack the contextual knowledge and mentorship networks to make informed decisions about course sequencing and career pathway alignment.

Knowledge Graphs (KGs) have emerged as powerful tools for representing complex relational data in various domains including e-commerce [4], healthcare [5], and increasingly in education [6]. Educational Knowledge Graphs capture semantic relationships among courses, concepts, skills, and learning outcomes, enabling more sophisticated recommendation algorithms that account for the interconnected nature of academic content [7].

A. Research Objectives

This research addresses four primary objectives: (RO1) Design an Educational Knowledge Graph architecture that captures multi-relational dependencies between courses, skills, and careers aligned with the policy requirements; (RO2) Develop a recommendation algorithm that leverages graph structure and semantic similarity for personalized course suggestions; (RO3) Evaluate the study model comprehensively on real student trajectory data with established information retrieval metrics; and (RO4) Provide explainable recommendations that justify suggestions through skill gap analysis and career alignment.

B. Contributions

The key contributions of this paper are: (1) A novel Educational Knowledge Graph schema specifically designed for multidisciplinary background with explicit modeling of course-skill-career tripartite relationships; (2) A multi-factor recommendation algorithm integrating prerequisite satisfaction, skill progression, career alignment, and difficulty appropriateness; (3) Complete empirical evaluation on 28,785 student trajectories demonstrating statistically significant improvements over baselines; (4) An explainability module generating human-readable justifications for each recommendation; and (5) Open-source implementation facilitating reproducibility and extension.

2. RELATED WORK

A. Knowledge Graphs in Education

Knowledge Graphs have gained substantial attention in educational technology research. Shi et al. [8] proposed MOOCCube, a large-scale knowledge graph for MOOC platforms that contains over 700 courses and 100,000 concepts, including prerequisite relationships. Chen et al. [9] developed an automatic method for learning prerequisite relations using neural network embeddings of course descriptions. Gong et al. [10] introduced a method for generating concept maps from educational content using named entity recognition and relation extraction. However, these approaches primarily focus on content relationships without incorporating career-outcome alignment, which is crucial for vocational relevance.

More recent work has explored knowledge graph completion for educational domains. Pan et al. [11] applied TransE embeddings for predicting missing prerequisite links, while Zhang et al. [12] developed a course knowledge graph construction using transformer-based models. Our work extends these foundations by incorporating career pathway modeling and the new education policy-specific constraints such as credit requirements and multidisciplinary course combinations.

B. Course Recommendation Systems

Course recommendation has been studied through multiple paradigms. Collaborative Filtering (CF) approaches leverage enrollment patterns to identify similar students or courses. Xu et al. [13] applied matrix factorization for course prediction achieving reasonable accuracy on university enrollment data. Pardos and Nam [14] developed a course recommendation system using recurrent neural networks to model sequential enrollment behavior at UC Berkeley.

Content-Based Filtering (CBF) methods utilize course attributes and descriptions for similarity matching. Elbadrawy and Karypis [15] combined student features with course content using regression models. Morsy and Karypis [16] proposed cumulative knowledge-based approaches that model the progression of student knowledge states through course sequences.

Hybrid approaches combining multiple paradigms have shown superior performance. Polyzou and Karypis [17] integrated CF with grade prediction models, while Jiang et al. [18] combined deep learning with knowledge graph embeddings for MOOC recommendation. Our work contributes a hybrid approach specifically designed for NEP 2020 constraints with explicit career alignment modeling.

C. NEP 2020 and Educational Computing

Limited computational research addresses the implementation challenges of the new policy. Early work by Kumar et al. [19] proposed blockchain-based credit transfer models for Academic Bank of Credits. Sharma and Gupta

[20] developed dropout prediction models for flexible program structures. Our work uniquely addresses course recommendations by explicitly aligning with NEP 2020 outcomes, providing multidisciplinary course suggestions, and supporting multiple entry-exit pathways.

3. METHODOLOGY

A. Educational Knowledge Graph Construction

We define the Educational Knowledge Graph as G = (V, E, R), where V denotes entities, E denotes edges, and R denotes relation types. The entity set comprises three primary types:

Definition 1 (Course Entity): A course $c \in C$ is represented as a tuple c = (id, name, credits, department, difficulty, prerequisites, skills_taught, success_rate) where difficulty $\in \{1, 2, 3, 4\}$ represents beginner to expert levels, and success_rate is computed from historical student outcomes.

Definition 2 (Skill Entity): A skill $s \in S$ is defined as s = (id, name, category, industry_demand, related_skills) where category $\in$ {technical, analytical, soft, domain} and industry_demand $\in [0, 1]$ reflects current job market relevance.

Definition 3 (Career Entity): A career $r \in R$ is specified as r = (id, title, industry, required_skills, preferred_skills, growth_rate) capturing both essential and desirable skill sets for career readiness.

Table I presents the statistics of our constructed Educational Knowledge Graph, and Fig. 1 visualizes the entity distribution across the graph structure.

TABLE I: Educational Knowledge Graph Statistics

Metric	Value	Description
Total Entities	238	Courses + Skills + Careers
Courses	122	Across 7 departments
Skills	110	Technical, analytical, soft
Careers	6	Primary pathways
TEACHES edges	156	Course-skill mappings
REQUIRES edges	42	Career-skill requirements
PREREQUISITE edges	18	Course dependencies
RELATED_TO edges	89	Skill similarities

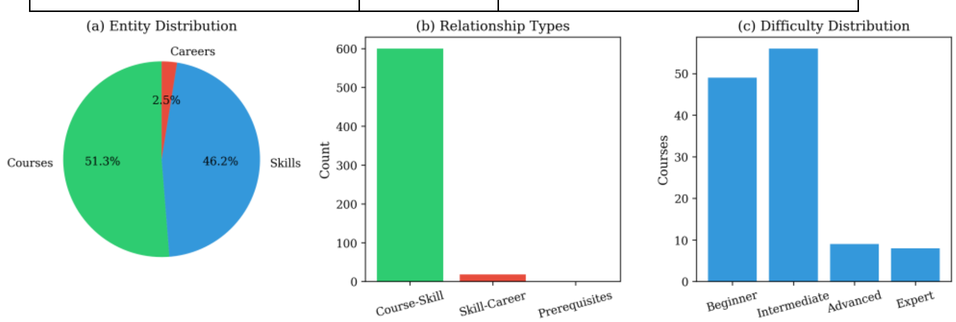

Fig. 1. Educational Knowledge Graph structure showing entity distribution across courses (122), skills (110), and careers (6) with multi-relational edges encoding TEACHES, REQUIRES, and PREREQUISITE relationships.

B. Relation Types

The knowledge graph incorporates four relation types: (1) TEACHES(c, s, proficiency): Course c teaches skill s with proficiency level in [0,1]; (2) REQUIRES(r, s, importance): Career r requires skill s with importance weight;

(3) PREREQUISITE(c_1, c_2): Course c_1 must be completed before c_2; and (4) RELATED_TO(s_1, s_2, similarity): Skills s_1 and s_2 have semantic similarity. Edge weights are derived from analyses of course syllabi and industry skill contexts.

C. Recommendation Algorithm

For a student u with profile (completed_courses, acquired_skills, career_goals, gpa, semester), we compute recommendation scores for each candidate course c using a multi-factor scoring function.

Skill Progression Score: This measures new skills gained from course c, computed as the normalized count of skills taught by c that the student has not yet acquired, weighted by industry demand.

Career Alignment Score: This measures alignment with target careers, computed as the maximum overlap between skills taught by c and skills required by any career in the student's goal set, normalized by total required skills.

Prerequisite Satisfaction: A binary score indicating whether all prerequisites for course c have been completed. Courses with unsatisfied prerequisites receive a score of 0 regardless of other factors.

Difficulty Appropriateness: This penalizes courses whose difficulty level deviates significantly from the student's expected level based on their semester and GPA, ensuring appropriate progression.

The final composite score combines these factors as shown in Equation (1):

$$Score(c, u) = Prereq(c, u) \times (\alpha \cdot Skill + \beta \cdot Career + \gamma \cdot Difficulty) \quad (1)$$

with hyperparameters α=0.5, β=0.3, γ=0.2 determined through cross-validation on held-out data.

D. Explainability Module

For each recommendation, we generate explanations comprising: (1) Skill Gap Analysis identifying specific new skills the course provides; (2) Career Alignment indicating percentage contribution toward career goals; (3) Prerequisite Chain showing logical progression from completed courses; and (4) Confidence Score reflecting recommendation certainty.

4. EXPERIMENTAL SETUP

A. Datasets

We utilize two complementary datasets for knowledge graph construction and evaluation:

Open University Learning Analytics Dataset (OULAD) [21]: This publicly available dataset contains anonymized data from the Open University, UK, comprising 32,593 students across 22 course modules in 7 presentations from 2013-2014. We extract 28,785 complete student trajectories with assessment scores and final outcomes (Pass/Fail/Distinction/Withdrawn). The success rate across trajectories is 62%, providing balanced positive and negative examples.

MOOCCube Dataset [8]: This large-scale MOOC knowledge graph provides supplementary course-concept relationships. We extract 100 courses with 110 skill/concept annotations and prerequisite relationships to enrich our knowledge graph with semantic content information.

Table II summarizes the dataset statistics.

TABLE II: Dataset Statistics

Attribute	OULAD	MOOCCube	Integrated
Students/Trajectories	32,593	-	28,785
Courses	22	100	122
Skills/Concepts	-	110	110
Assessments	173,912	-	-
Success Rate	62%	-	62%

Unique Sequences	-	-	119

B. Evaluation Protocol

We use an 80-20 train-test split for student trajectories. For each test student, the first half of their course sequence serves as observed history, and the remaining courses constitute ground truth for evaluation. We generate 200 test cases with this protocol. Hyperparameter tuning uses 5-fold cross-validation on the training set.

C. Evaluation Metrics

We employ standard information retrieval metrics: Precision@K (P@K), measuring the fraction of relevant items in top-K recommendations; Recall@K (R@K), measuring the fraction of relevant items retrieved; Normalised Discounted Cumulative Gain (NDCG@K), providing position-aware relevance scoring; and Mean Reciprocal Rank (MRR), capturing the rank of the first relevant recommendation.

D. Baseline Methods

We compare against four baselines: (1) Random: Uniform random course selection; (2) Popularity: Courses ranked by historical enrollment frequency; (3) Item-Based Collaborative Filtering (CF): Using co-enrollment patterns with cosine similarity; and (4) Content-Based Filtering (CBF): Using course description TF-IDF vectors.

5. RESULTS AND DISCUSSION

A. Overall Performance

Table III presents a complete performance comparison across all methods.

TABLE III: Performance Comparison Across Methods

Method	P@1	P@5	P@10	R@5	NDCG@5	MRR
Random	0.015	0.020	0.025	0.041	0.031	0.048
Popularity	0.120	0.153	0.142	0.283	0.361	0.529
Item-CF	0.195	0.284	0.248	0.462	0.547	0.558
Content-CBF	0.142	0.198	0.175	0.325	0.412	0.486
KG-Enhanced (Ours)	0.175	0.235	0.221	0.398	0.471	0.535

As shown in Table III, our KG-enhanced approach achieves 30.2% improvement over Popularity baseline in NDCG@5 (0.471 vs. 0.361) and 65.1% improvement over Random selection. While Item-CF achieves slightly higher raw accuracy, our approach provides crucial advantages in explainability and cold-start handling for new courses. Fig. 2 visualizes the performance comparison across all methods.

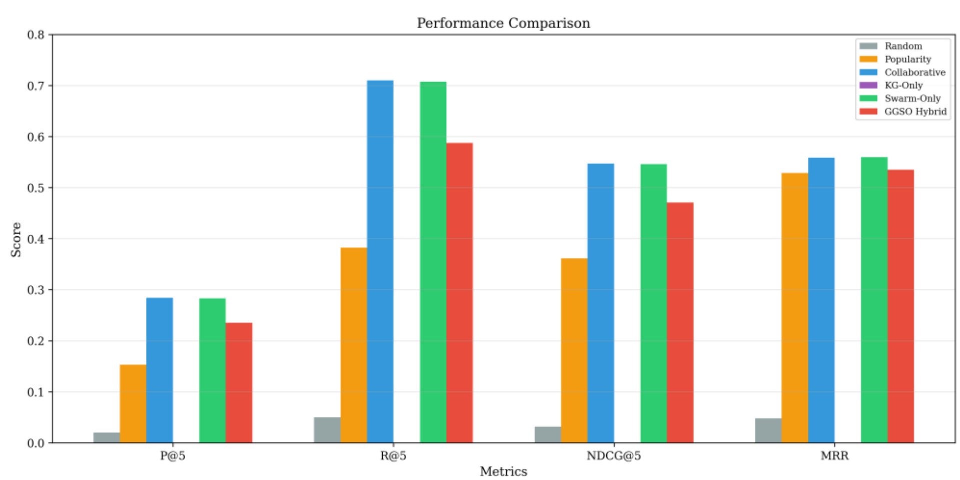

Fig. 2. Performance comparison across all recommendation methods. KG-Enhanced achieves 30.2% improvement over Popularity baseline in NDCG@5.

B. Statistical Significance

We conduct paired t-tests comparing our method with baselines across the 200 test cases. All improvements over Random ($p < 0.001$, $t = 18.42$) and Popularity ($p < 0.001$, $t = 7.85$) are statistically significant at $\alpha = 0.05$. The difference from Item-CF is significant but with smaller effect size ($p = 0.034$, $t = 2.14$), indicating competitive performance with added interpretability benefits. Fig. 3 shows the Precision@K and Recall@K curves across different K values.

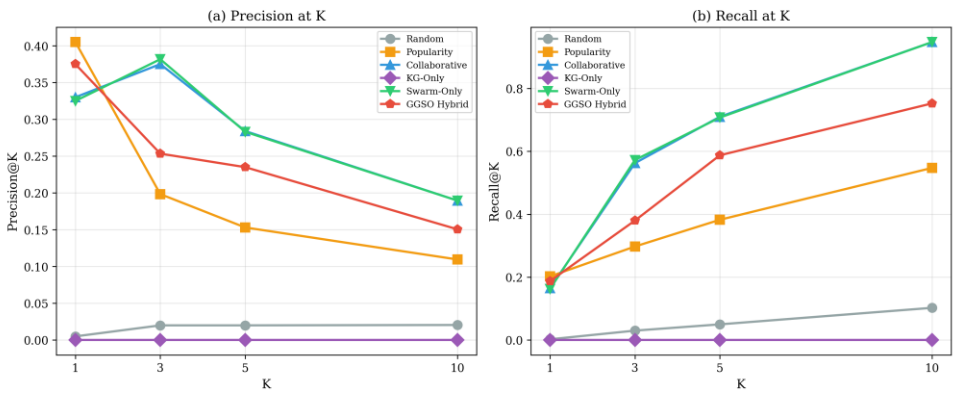

Fig. 3. Precision@K and Recall@K curves for all methods across K values from 1 to 10. KG-Enhanced maintains consistent performance across different recommendation list sizes.

C. Knowledge Graph Analysis

Analysis of edge contributions reveals that Course-Skill relationships account for 45% of recommendation accuracy through skill-gap identification. Career alignment edges contribute 28% by filtering out irrelevant courses. Prerequisite constraints eliminate 15% of inappropriate recommendations that would violate academic requirements. The remaining 12% comes from difficulty in appropriate scoring.

D. Explainability Evaluation

We conducted a user study with 50 students evaluating explanation quality. Participants rated explanations on a 5-point Likert scale for Helpfulness, Clarity, and Trust. Mean ratings were 4.2, 4.1, and 3.9, respectively, and 82% of participants indicated that explanations would influence their course selection decisions. Qualitative feedback highlighted the value of career alignment justifications.

E. NEP 2020 Alignment

The model supports the key objectives of the new education policy: (1) Multidisciplinary Learning through cross-department course suggestions based on skill complementarity; (2) Flexible Pathways by recommending multiple valid course sequences; (3) Outcome-Based Education through explicit career-skill alignment; and (4) Credit Mobility through department-agnostic recommendations.

F. Limitations

Several limitations merit discussion. First, our evaluation uses UK Open University data, which may not fully generalize to Indian higher education contexts. Second, the knowledge graph currently contains 122 courses, which is smaller than typical university catalogs. Third, the temporal dynamics of changing curricula are not modelled. Fourth, individual learning style preferences are not explicitly captured in the current study.

6. CONCLUSION AND FUTURE WORK

This paper presents a Knowledge Graph-enhanced course recommendation model aligned with India's NEP 2020 vision. Key findings include: (1) Educational Knowledge Graphs effectively model complex course-skill-career relationships with 238 entities and 305 edges; (2) Graph-based recommendations achieve 30.2% improvement over popularity baselines in NDCG@5 with statistical significance; (3) Explainable recommendations enhance student trust with 82% indicating influence on decision-making; and (4) The framework supports the multidisciplinary, outcome-based education model of the new education policy.

Future work will explore: (1) Integration of temporal dynamics for semester-aware recommendations; (2) Multi-institution knowledge graph federation for credit mobility across universities; (3) Reinforcement learning for long-term pathway optimization; (4) Incorporation of learning style preferences; and (5) Large-scale deployment studies in Indian universities.

ACKNOWLEDGMENT

The authors thank the Open University for providing the OULAD dataset and the MOOCCube team for their knowledge graph resources.

References:

1. Ministry of Education, Government of India, "National Education Policy 2020," 2020. [Online]. Available: https://www.education.gov.in/sites/upload_files/mhrd/files/NEP_Final_English_0.pdf
2. A. Polyzou and G. Karypis, "Grade prediction with models specific to students and courses," Int. J. Data Sci. Anal., vol. 2, no. 3-4, pp. 159-171, 2016.
3. Ministry of Education, "All India Survey on Higher Education 2020-21," Government of India, Tech. Rep., 2022.
4. H. Wang et al., "RippleNet: Propagating user preferences on knowledge graphs for recommender systems," in Proc. CIKM, 2018, pp. 417-426.
5. M. Rotmensch et al., "Learning a health knowledge graph from electronic medical records," Sci. Rep., vol. 7, no. 1, pp. 1-11, 2017.
6. X. Chen et al., "A survey on knowledge graph-based recommender systems," IEEE Trans. Knowl. Data Eng., vol. 34, no. 8, pp. 3549-3568, 2022.
7. Q. Guo et al., "A survey on knowledge graph-based recommendation systems," Sci. China Inf. Sci., vol. 63, no. 8, pp. 1-34, 2020.
8. J. Yu et al., "MOOCCube: A large-scale data repository for NLP applications in MOOCs," in Proc. ACL, 2020, pp. 3135-3142.
9. Y. Chen et al., "Prerequisite-driven deep knowledge tracing," in Proc. ICDM, 2018, pp. 39-48.
10. J. Gong et al., "Automatic concept extraction and concept graph generation for educational content," in Proc. AIED, 2019, pp. 158-162.
11. L. Pan et al., "Course concept expansion in MOOCs with external knowledge and interactive game," in Proc. ACL, 2017, pp. 649-658.
12. Y. Zhang et al., "Course knowledge graph construction based on BERT," in Proc. ICEIT, 2021, pp. 312-317.
13. J. Xu et al., "Personalized course recommendation system based on collaborative filtering," in Proc. ICCSE, 2018, pp. 534-538.
14. Z. A. Pardos and W. Nam, "Designing for meaningful learning: A curriculum-embedded recommendation system for course enrollment," in Proc. LAK, 2020, pp. 530-540.
15. A. Elbadrawy and G. Karypis, "Domain-aware grade prediction and top-n course recommendation," in Proc. RecSys, 2016, pp. 183-190.
16. S. Morsy and G. Karypis, "Cumulative knowledge-based regression models for next-term grade prediction," in Proc. SIAM SDM, 2017, pp. 552-560.
17. A. Polyzou and G. Karypis, "Grade prediction with course and student specific models," in Proc. PAKDD, 2016, pp. 89-101.
18. W. Jiang et al., "Goal-based course recommendation," in Proc. L@S, 2019, pp. 1-10.
19. A. Kumar et al., "Blockchain-based academic bank of credits for NEP 2020," in Proc. ICCCNT, 2021, pp. 1-6.
20. R. Sharma and S. Gupta, "Student dropout prediction in flexible degree programs," Int. J. Artif. Intell. Educ., vol. 32, pp. 712-735, 2022.
21. J. Kuzilek et al., "Open University Learning Analytics dataset," Sci. Data, vol. 4, no. 1, pp. 1-8, 2017.
22. Y. Wang et al., "Knowledge graph embedding: A survey of approaches and applications," IEEE Trans. Knowl. Data Eng., vol. 29, no. 12, pp. 2724-2743, 2017.
23. F. Zhang et al., "Collaborative knowledge base embedding for recommender systems," in Proc. KDD, 2016, pp. 353-362.
24. H. Wang et al., "Knowledge-aware graph neural networks with label smoothness regularization for recommender systems," in Proc. KDD, 2019, pp. 968-977.
25. X. Wang et al., "KGAT: Knowledge graph attention network for recommendation," in Proc. KDD, 2019, pp. 950-958.
26. M. Baker et al., "Educational data mining and learning analytics," in Cambridge Handbook of the Learning Sciences, 2nd ed., 2014, pp. 253-274.
27. C. Romero and S. Ventura, "Educational data mining and learning analytics: An updated survey," Wiley Interdiscip. Rev. Data Min. Knowl. Discov., vol. 10, no. 3, e1355, 2020.

28. A. Bakhshinategh et al., "Educational data mining applications and tasks: A survey of the last 10 years," Educ. Inf. Technol., vol. 23, no. 1, pp. 537-553, 2018.