Svedberg Open
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

Integrating Knowledge Graphs with Natural Language Processing for Context-Aware Educational Content Recommendations

Dr. Priya Sethuraman[1*], Dr. Arivukkodi R[2], Dr. Nallusamy C[3], Durga B[4], Xalida Sultanova[5], Bakhodir Khoshbakov[6]

[1*]Professor, Department of Management Studies, St. Joseph's Institute of Technology, OMR, Chennai, Tamil Nadu, India. Email: priasethuraman@gmail.com

[2]Assistant Professor, Computer Science, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, Chennai, Tamil Nadu, India. Email id: arivukodir@maher.ac.in

[3]Professor, Department of Information Technology, K.S. Rangasamy College of Technology, Tiruchengode, Namakal, Tamil Nadu, India. Email: nallusamyc@ksrct.ac.in, https://orcid.org/0000-0001-6100-0088

[4]Associate Professor, Meenakshi College of Allied Health Sciences, Meenakshi Medical College Hospital & Research Institute, Meenakshi Academy of Higher Education and Research, Chennai, Tamil Nadu, India. Email id: durgab@maher.ac.in

[5]Department of Social Sciences, Tashkent State Medical University Tashkent, Uzbekistan. E-mail: sultanovaholida9@gmail.ru, https://orcid.org/0000-0003-4909-2121

[6]Department of Economics, Termez University of Economics and Service
Termez, Uzbekistan. E-mail: baxodir_xushbakov@tues.uz, https://orcid.org/0009-0006-7566-6976

*Corresponding author: Email: priasethuraman@gmail.com

Abstract

Personalization of content recommendation for education has resulted in a combination of knowledge graphs and natural language processing in order to develop context-aware recommendation systems. The proposed paper presents the KG-NLP-CAR (Knowledge Graph and Natural Language Processing for Context-Aware Recommendations) model, that allows for dynamic personalization of recommendations based on the power of reasoning inherent in knowledge graphs and the contextualization provided by natural language processing technologies. Specifically, the uniqueness of the proposed model is associated with using specific information about each individual learner such as preferences, goals, and interactions with the platform, all of which are represented within a knowledge graph. Preprocessing of the educational content is performed using NLP tools such as tokenization, entity recognition, and generation of semantic embeddings. Thus, the model is able to give recommendations for content that will be most useful for the user according to his context. The performance of the model was evaluated on Last.FM, Book-Crossing, and MovieLens-1M datasets, where the KG-NLP-CAR model performed better than the several state-of-the-art models. It gave a very high AUC value (95.65%) and an F1 score of 88.45% when evaluated on the MovieLens-1M dataset, easily beating Ripple Net and KGAT in all these categories. The ablation studies conducted showed that KGs and NLP were key factors in making the recommendations accurate and relevant. Thus, the KG-NLP-CAR model shows that the combination of KGs and NLP can be effectively used to give educational recommendations.

Keywords

Knowledge Graphs, Natural Language Processing, Context-Aware Recommendations, Personalized Learning, Educational Content, Embedding Generation, Performance Evaluation

1. Introduction

For the integration of KGs and NLP for educational content recommendations in the context of knowledge graphs, the hybrid model utilizes the reasoning capabilities of the knowledge graph and the flexibility of natural language processing to make accurate recommendations. The model enables the extraction and understanding of educational contents based on the learners' preferences, goals, and context. Preprocessing of the educational contents is done using the capabilities of natural language processing. These include tokenization, named entity recognition, and contextual embedding. The research focuses on context-aware recommendation systems, specifically those involving knowledge graphs and NLP. Contextual knowledge graphs and NLP can be used in the educational context to create personalized content recommendations, hence being highly applicable for the integration of KGs and NLP for educational content recommendations [1]. The research describes an improved knowledge recommendation method for manufacturing process planning using context-aware knowledge graphs. The methodology described can be adopted in educational contexts through modification of the context and representation of knowledge with Numerical methods [2].

The presented paper is focused on presenting a new transfer deep learning approach with cross-lingual embeddings applied to sentiment analysis. The discussed model could be useful for educational content recommendation systems as it increases the performance of NLP techniques related to personalized content retrieval[3].Their research presents an example of combining knowledge graphs and contextual information for effective application in educational recommendations[4].In this research, the main emphasis was put on personalizing educational content recommendation approaches by means of cosine similarity-based knowledge graphs and contextual awareness. Thus, such an aspect correlates well with the improvement of educational content recommendation with NLP and KGs [5][23]. The article presents the usage of knowledge management systems at higher education institutions and the importance of using structured representation of knowledge for this purpose. In this case, the usage of KGs for educational recommendations could be viewed through the prism of that framework with NLP support. The paper proposes the use of a context-aware collaborative recommendation system based on knowledge graphs. It highlights the significance of implementing collaborative filtering, together with context information, which can be applied to recommending educational content that incorporates KGs and NLP [7][21]. This systematic review examines recent developments in context-aware recommendation systems. It highlights how combining KGs with NLP enhances recommendation systems, thereby providing a good basis for developing personalized educational content using a similar hybrid approach [8][22]. The article illustrates the benefits of integrating KGs with NLP for delivering personalized educational content. Specifically, it captures learner contexts while delivering personalized educational content [9]. The study proposes an AI-based content recommendation system in LMS using hybrid filtering. It shows how incorporating KGs with NLP can help enhance the effectiveness of the hybrid recommendation model in educational environments [10][24]. The research examines a knowledge-aware recommendation system and AI based Mathematical approach using enhanced contrastive learning, with a focus on integrating knowledge graphs and NLP. It demonstrates the use of contextual knowledge representation for improving educational content recommendations [11][19].

Key Contribution

1. This paper proposes a KG-NLP-CAR model that integrates Knowledge Graph (KG) and Natural Language Processing (NLP) for providing personalized education content recommendations according to their preference and needs.
2. For providing such personalized recommendations, the proposed model employs context-aware strategies by considering the learners' current state to make them more effective.
3. The performance of the KG-NLP-CAR model is better than other models, and it provides very accurate results as well.

This research is followed by the various sections. Section I introduces the research, Section II provides a literature review of the work, and Section III explains the proposed methodology for the overall architecture and the Knowledge Graph for Context-Aware Educational Context Recommendations. Section IV explained the results and discussion, followed by a dataset description and parameter initialization with various analyses. Section V explained the conclusion of this research.

2. Literature Review

Table1: Summary of Related Work

Reference	Model Name	Applications	Key Findings
[12]	LLM-driven context-aware recommendation system	Social media personalization, educational content recommendations	The model enhances personalization by integrating NLP with context-awareness, improving content recommendation accuracy and relevance, particularly in dynamic social contexts like education.
[13]	None (the study is conceptual)	Language learning, EFL vocabulary enhancement, educational content recommendation	Explores the relationship between vocabulary depth and writing ability, suggesting that better vocabulary understanding can lead to improved writing skills, which can inform personalized educational content based on learner vocabulary proficiency.
[14]	Smart context-aware learning system	English language learning, writing, speaking enhancement	The study shows that context-aware systems significantly improve writing and speaking skills by adapting to learner needs, suggesting that similar context-aware models could be applied to educational content recommendations in personalized learning systems.
[15]	Bidirectional LSTM	Sentiment classification, educational content sentiment analysis	The model uses bidirectional LSTM to classify sentiments in movie reviews. This approach can be adapted for context-aware educational content recommendations that require sentiment analysis for personalized content delivery.
[16]	Explainable IR techniques	Academic search engines, personalized content retrieval	The use of explainable IR techniques enhances content retrieval in academic contexts, making it easier to adapt these methods for educational recommendation systems that require interpretability and transparency.
[17]	Adaptive web-based search framework	Web-based educational search, content personalization	Integrating NLP and ML into web-based frameworks improves search relevance and content personalization, helping to adapt recommendations to learners' individual needs and learning goals.
[18]	Context-aware recommendation framework	Personalized content recommendation, user clustering, sentiment analysis	The hybrid framework combining user clustering with BERT-based sentiment analysis improves recommendation accuracy by better aligning content with learner sentiment and preferences.
[6]	Knowledge graph learning model	Point-of-interest recommendation, educational content personalization based on learner interest	The model uses context-aware knowledge graph learning to enhance recommendations by considering both context and user preferences, which can be adapted for personalized educational content recommendations based on learner interest and context.
[20]	Knowledge graph-integrated recommendation method	Career planning, personalized educational recommendations for students	The knowledge graph-based recommendation method aids in career planning by integrating academic data with personalized suggestions, offering a model that can also be applied to educational content recommendations based on student goals.

Table1 summarizes the different models available for context-aware recommendations of educational content. Various methods, including NLP, machine learning, and knowledge graph, have been combined in these models to achieve personalized content delivery. For example, the LLM context-aware recommendation model enhances the accuracy of content by applying NLP to it. Vocabulary enhancement research for English as a Foreign Language indicates that the depth of vocabulary is related to writing ability and may help personalize educational content. A smart context-aware learning system adapts to the learner's needs, thus enhancing writing and speaking abilities. The bidirectional LSTM model and the explainable information retrieval models emphasize the importance of sentiment analysis and transparent information retrieval, respectively, in tailoring educational material. The adaptive web-based search model relies on NLP and machine learning algorithms to provide content that is more relevant to the learner, whereas the context-aware recommendation framework and knowledge graph learning models take into account sentiments, preferences, and interests to refine the recommendations made. Finally, the recommendation model based on knowledge graphs aids in career and academic goal planning, offering tailored educational content suggestions to learners.

Research Gap

An important research gap in Integrating Knowledge Graphs with Natural Language Processing for Context-Aware Educational Content Recommendations includes gaining insight into how to effectively modify content recommendations in real time to adapt to changes in the context of learning. Although existing models demonstrate good performance when applying static personalization, it would be necessary to develop a system capable of adjusting itself according to the dynamics of learners' advancement, emotional state, and other environmental conditions. Moreover, the integration of various learning styles and social contexts into the knowledge graph is yet to be fully explored, reducing the potential effectiveness of the model in recommending personalized education content.

3. Proposed Methodology

3.1 Overall Architecture of Knowledge Graph with Natural Language Processing for Context-Aware Educational Content Recommendations

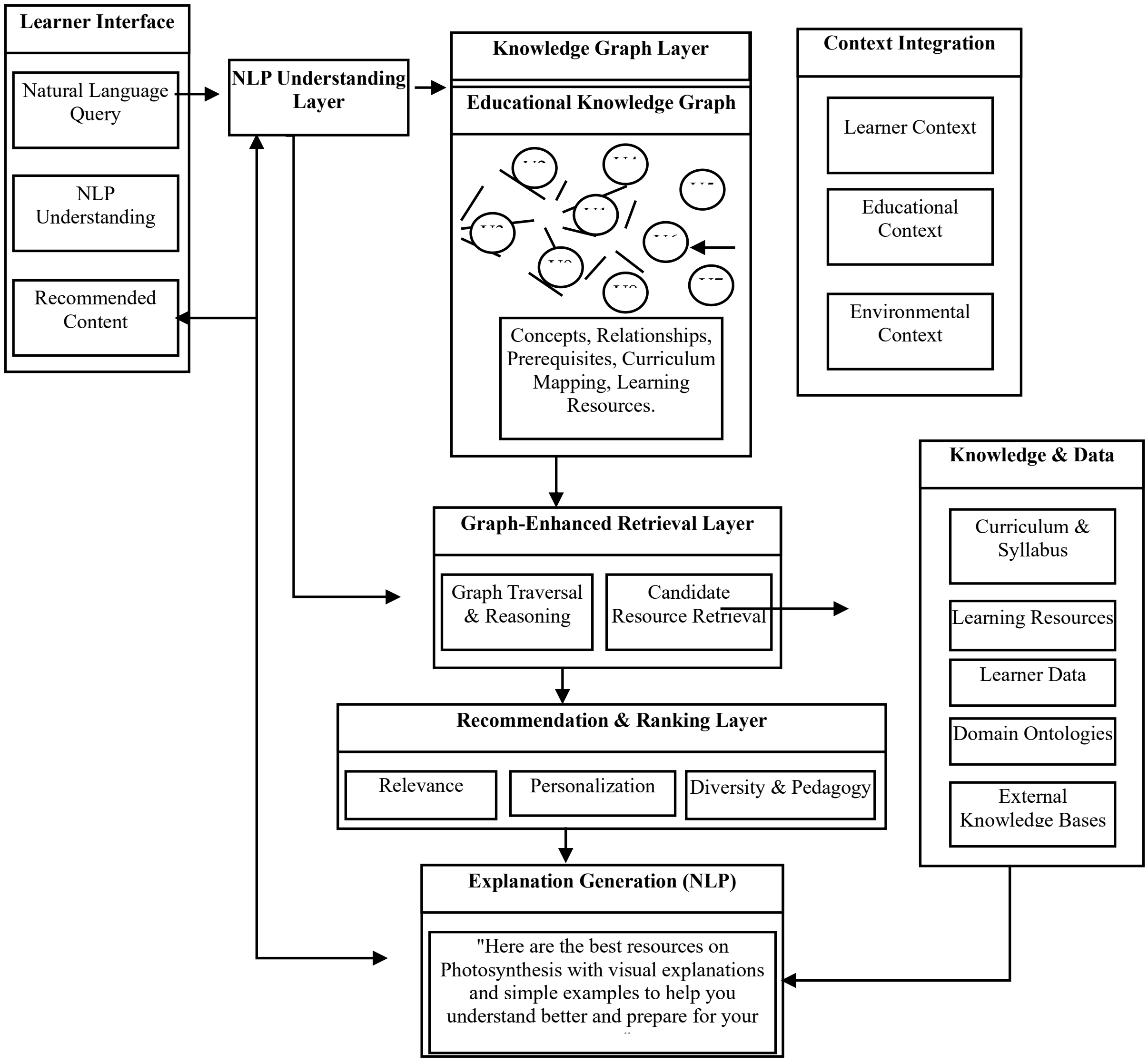

Figure 1: Overall Architecture for proposed Methodology

Fig1 describes the complex system that makes use of NLP and Knowledge Graphs in order to recommend personalized and contextual content of education. Then, it takes into account Knowledge Graph Layer that organizes all the educational concepts, relationships, and prerequisites in the form of an educational knowledge graph. This helps in identifying the relationships between various topics and ensures that the content recommended to the learner aligns with the educational goal of the user. The value of the request provided by the learner is enhanced by the Context Integration Layer in that the context of the learner such as knowledge level, learning goals, previous activities, time, and location of the learner is considered. Graph-Enhanced Retrieval Layer analyzes the knowledge graph in order to discover relevant concepts, identify associated topics, and verify prerequisites while expanding the knowledge neighborhood with more relevant items. Then the resources that satisfy the criteria of the learner's query are fetched and further refined based on the context of the learner. After this, the recommendation and ranking of these resources is done by the Recommendation & Ranking Layer based on relevance, difficulty, preferences of the learner, and the desired educational goals.Finally, the Explanation Generation (NLP) layer generates explanations for the recommended learning content. For instance, explanations about photosynthesis are included, along with diagrams and simple examples, to help the learner understand the subject matter and prepare for the exam. Robust technological foundations, including Knowledge Graphs (Neo4j/Janus Graph), NLP Models (BERT/RoBERTa), Vector Databases (Pinecone/FAISS), Generative AI models like GPT-4, and system performance analytics, underpin the whole system. This multi-layered structure ensures that the educational material is relevant to the learner's requirements and situation.

3.2 Knowledge Graph for Context-Aware Educational Content Recommendations

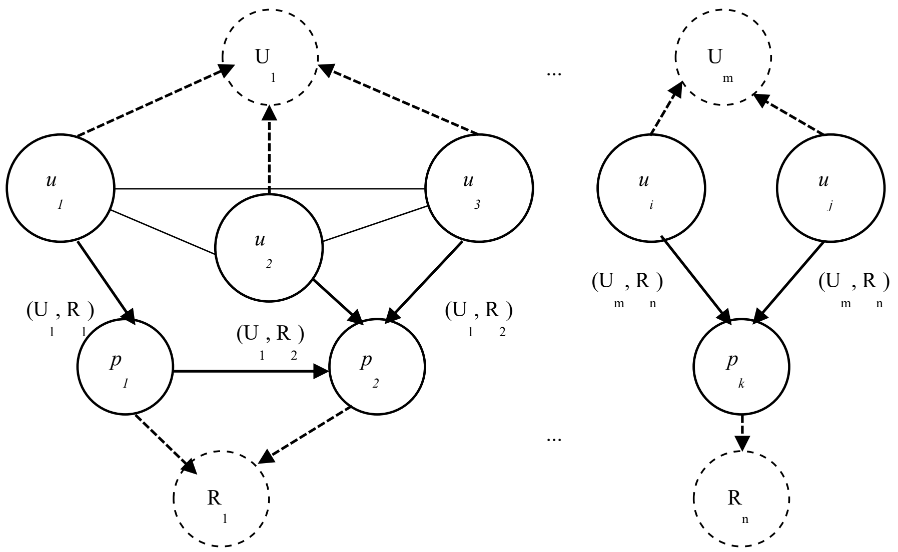

Figure 2: Knowledge Graph for Context-Aware Educational context Recommendations

Fig2 represents the knowledge graph schema used in the development of context-aware educational content recommendations. In particular, it provides an overview of the connections between distinct objects or entities. At the beginning of the graph, one can notice that there are several learners/users (blue nodes) who interact with several items of educational content (orange nodes). The educational content items are tagged with "p." In this case, the interaction is captured within a knowledge base. In other words, the knowledge base contains a collection of data describing the interactions of learners/users with the educational content items in question. Besides, the context such as educational course and learning level are also taken into account. The knowledge graph shows the way of interaction from learner/user to educational content item and their relations with specific resources referred to as "R." Moreover, the right part of the knowledge graph continues the idea. That is, the connections of multiple users (U) with different resources (R) in a generalized form.

In this defined the triplet should represent the (h, r, t) where h and t should represent the head and tail entities. r should represent the semantic relation. The main idea of this is to introduce the independent projection matrix that represents,

$$h_r = M_r h_r t_r = M_r t \tag{1}$$

From the above Equation (1) describes the h_r and t_r denotes the projection among the vectors of entities of related specific space and M_r should represents the projection matrix associated with the relation of r.

During the embedding training process, the model should represent the margin based on the employed to improve the model's ability to discriminate between the positive and negative triplets. The scoring functions are defined as the positive and negative samples are defined as,

$$f_r(h, t) = \| h_r + r_{s-t} - t \|_2^2 \tag{2}$$

From the above Equation (2) represents the $f_r(h, t)$ represents the score of the function based on positive triplet and h_r, t_r and r_{s-t} defined the embedding vectors of the head entity, tail entitiy and semantic relation.

3.3 Natural Language Processing for Context-Aware Educational Content Recommendations

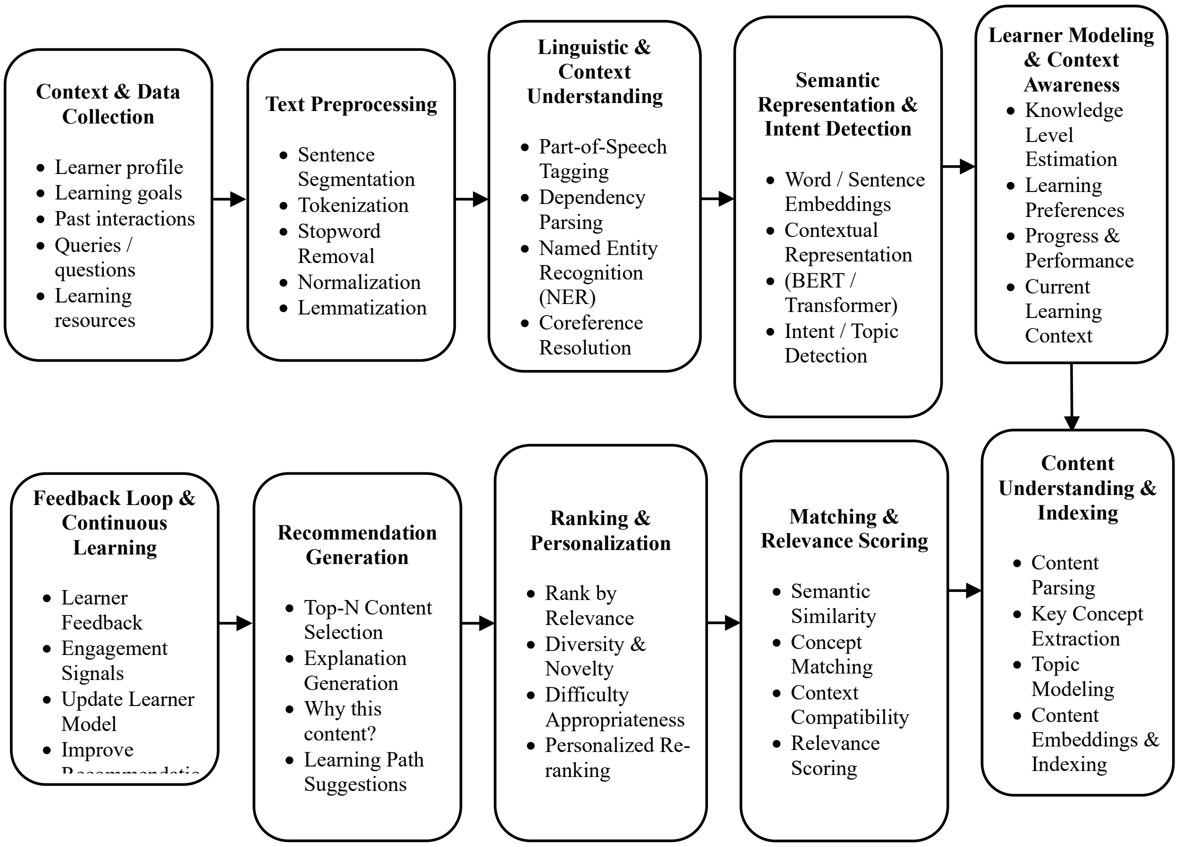

Figure 3: Natural Language processing for context Aware Educational Content Recommendations

Fig3 shows the Comprehensive pipeline for Natural Language Processing (NLP) aimed at delivering context-aware educational content recommendations. It begins with Context & Data Collection, which gathers essential learner data such as profiles, goals, past interactions, queries, and learning resources. The collected data undergoes Text Preprocessing, involving processes like sentence segmentation, tokenization, stop word removal, and lemmatization to prepare the text for further analysis. The next step, Linguistic & Context Understanding, includes part-of-speech tagging, dependency parsing, and named entity recognition (NER) to gain a deeper understanding of the text's structure and meaning. Following this, Semantic Representation & Intent Detection uses word and sentence embeddings, contextual representation through models like BERT, and intent/topic detection to capture the semantic nuances of the content. The pipeline continues with Learner Modeling & Context Awareness, where the system assesses the learner's knowledge level, learning preferences, progress, and current learning context. Content Understanding & Indexing ensures that educational materials are parsed, indexed, and categorized by key concepts, making them easily retrievable for learners. Matching & Relevance Scoring then evaluates content based on its semantic similarity, concept matching, and context compatibility, ensuring the right content is aligned with the learner's needs. In the Ranking & Personalization phase, content is ranked by relevance, diversity, novelty, and appropriateness to the learner's difficulty level, ensuring personalized re-ranking for optimal learning experiences. Finally, a Feedback Loop & Continuous Learning is implemented to refine the learner model based on engagement signals and feedback, improving content recommendations over time. The entire process ensures that the right educational content is delivered to the right learner, in the right context, at the right time, enhancing engagement and performance while providing a personalized learning journey.

3.4 Proposed Algorithm

Algorithm: Integrating Knowledge Graphs with NLP for Educational Content Recommendations

Input:

- Educational content (e.g., text-based materials)
- Learner profile (including preferences, learning history, and goals)

Output:

- Personalized educational content recommendations

Step 1: Preprocess Educational Content

1. Tokenize the educational content into words or phrases.
2. Extract Entities from the content using NLP techniques (e.g., Named Entity Recognition).
3. Generate Embeddings for the content to represent its semantic meaning using NLP models like BERT or GPT.

Step 2: Build/Update the Knowledge Graph

4. For each extracted entity, check if it already exists in the Knowledge Graph:

○ If it doesn't exist, add the entity to the Knowledge Graph.

5. For each embedding, find semantically similar entities and create relationships between them in the Knowledge Graph.

Step 3: Personalize Recommendations

6. Retrieve the learner's context (e.g., learning goals, preferences).
7. Query the Knowledge Graph for relevant entities based on the learner's context.
8. Rank the content based on relationships in the graph and the learner's profile.

Step 4: Return Recommended Content

9. Return a list of recommended educational content based on the ranked entities and their relationships in the Knowledge Graph.

End of Algorithm

The proposed algorithm integrates Knowledge Graphs (KGs) and Natural Language Processing (NLP) to deliver personalized educational content recommendations. First, the algorithm preprocesses the content by tokenizing it, extracting entities using NLP techniques like Named Entity Recognition, and generating semantic embeddings through models like BERT or GPT. It then updates the Knowledge Graph by adding new entities and establishing relationships between semantically similar entities. To personalize the recommendations, the learner's profile, including their goals and preferences, is used to query the Knowledge Graph for relevant content, which is then ranked based on its contextual relevance to the learner. Finally, the algorithm returns a list of personalized educational content, ensuring that the recommendations align with the learner's needs and learning objectives, ultimately enhancing their educational experience.

4. Results and Discussion

4.1 Dataset Description

Last.FM, Book-Crossing, and MovieLens-1M are some of the useful datasets that could be used to integrate Knowledge Graphs (KGs) and Natural Language Processing (NLP) for context-aware educational content recommendations. For instance, Last.FM dataset includes users' listening habits, tracks' information, artist information, album information, and user tags, thus making it suitable for developing knowledge graphs that will show relationships between users, tracks, genres, and tags. Additionally, using NLP, it would be possible to analyze user generated tags such as relaxing and upbeat among others, and develop context-aware recommendation engine. In the same way, Book-Crossing dataset, which consists of ratings and metadata information about the books, could also be used in constructing KGs with users, books, and genres among other things, and analyzing books' titles/descriptions using NLP. This allows personalization of books based on the user's rating history and context-based preferences. The MovieLens-1M dataset has comprehensive information about movies in terms of their ratings, user data, and other metadata like genres and description, which could be used to create a Knowledge Graph connecting users with movies, genres, and ratings. NLP techniques may be used to identify themes in the movie's description, enabling more accurate recommendations considering the user demographic and past interests. Through the use of the above data and models, personalized recommendations could be provided, taking into account not only the user's previous activity but also contextual factors like mood, time, and genre preference.

The TF, which stands for the TREC Fairness Dataset, belongs to the Text Retrieval Conference (TREC). This particular dataset is related to fairness in machine learning, more precisely, in question-answering systems; the main idea is that it offers data to be used when measuring and improving fairness in machine learning predictions. The other one is JD, which stands for Job Descriptions Dataset. This resource includes job postings and descriptions from multiple industries; it is used in job title classification, job skill extraction, and matching jobs with potential candidates. Finally, there is the TREC, or the Text Retrieval Conference Dataset. It is commonly known and used in information retrieval studies. Among the labeled data offered in this resource, there are labels for different tasks, including question classification, sentiment analysis, and document ranking.

Table 2: Description of Dataset Details

Datasets	Last. FM	Book-Crossing	movieLen-1M
Users	1872	17860	6036
Items	3864	14910	2445
Interactions	42346	139746	753772
Entities (KG)	9366	77903	182011
Relations (KG)	60	25	12
Triples (KG)[19]	15518	151500	1241996

Table2 shows the Last.FM, Book-Crossing, and movieLen-1M datasets consist of comprehensive details about their structure. The Last.FM dataset comprises 1,872 users, 3,864 items, and 42,346 interactions. Additionally, it has 9,366 entities and 60 relations in its Knowledge Graph (KG) that is made up of 15,518 triples. In comparison, the Book-Crossing dataset has 17,860 users, 14,910 items, and a relatively high number of interactions of 139,746. The Knowledge Graph (KG) of Book-Crossing includes 77,903 entities, 25 relations, and 151,500 triples. On the other hand, the movieLen-1M dataset involves 6,036 users and 2,445 items and a total of 753,772 interactions. Besides, it has 182,011 entities.

Table 3: Dataset Description about Recommendation System for Education

Data Description	TF	JD	TREC
Students	1409	6039	20860
Knowledge of career planning	5637	2445	15947

Interactions	26975	123772	162036
Entitles	8263	6011	10398
Relations	59	11	21
KG triples[20]	10552	136928	52379

Table3 gives us all relevant information related to three datasets which have been used for recommendation systems research purposes. First one is the Last.FM dataset which has 1,872 users, 3,864 items, and 42,346 interactions. In addition, the KG of Last.FM includes 9,366 entities and 60 relations with 15,518 triples. Secondly, Book-Crossing dataset includes 17,860 users, 14,910 items, and a high number of interactions (139,746). Book-Crossing KG has 77,903 entities, 25 relations, and 151,500 triples. The third dataset considered by us is the movieLen-1M dataset which has 6,036 users and 2,445 items with 753,772 interactions. Its KG has 182,011 entities, 12 relations, and 1,241,996 triples.

4.2 Software and Hardware Configurations

Table 4: Software and Hardware Configurations

Component	Specification
Processor	Intel Core i7-12700H, 2.30 GHz, 14-core
RAM	32 GB DDR5 @ 4800 MHz
GPU	NVIDIA GeForce RTX 3060 (6 GB GDDR6)
Operating System	Ubuntu 22.04 LTS (64-bit)
Programming Language	Python 3.10.6
Deep Learning Framework	TensorFlow 2.12.0 / Keras 2.12.0
ML Libraries	scikit-learn 1.2.2, imbalanced-learn 0.10.1
Data Processing	NumPy 1.24.3, Pandas 2.0.1
Visualization	Matplotlib 3.7.1, Seaborn 0.12.2
IDE / Environment	Jupyter Notebook 7.0.0 / VS Code 1.79

Table4 shows the Hardware and software requirements to integrate Knowledge Graphs with Natural Language Processing for context-aware educational content recommendations have been selected so that can provide maximum performance for deep learning and computing operations. It includes a computer with Intel Core i7-12700H having 14 cores and providing great processing power and 32 GB of DDR5 RAM to process data. It is coupled with the NVIDIA GeForce RTX 3060 graphics card having 6 GB of VRAM to accelerate deep learning computations for natural language processing. All the hardware components have been configured under Ubuntu 22.04 LTS version. The programming language being used is mainly Python version 3.10.6, while TensorFlow version 2.12.0 and Keras version 2.12.0 are utilized for training models using neural networks and deploying these models respectively. For the machine learning process, scikit-learn version 1.2.2 and imbalanced-learn version 0.10.1 are used for evaluating models and dealing with class imbalances respectively. NumPy version 1.24.3 and Panda's version 2.0.1 are used to manipulate and preprocess data, respectively, while Matplotlib version 3.7.1 and Seaborn version 0.12.2 are used to visualize results. Jupyter Notebook version 7.0.0 is used for writing code interactively, while VS Code version 1.79 is used for coding and debugging.

4.3 Parameter Initialization

Table5: Parameter Initialization

Parameter	Value
Learning Rate	0.001
Batch Size	32
Epochs	50
Optimizer	Adam
Loss Function	Sparse Categorical Cross-Entropy
Embedding Dimension	128
Dropout Rate	0.3
Activation Function	ReLU

Hidden Layers	[128, 64, 32]
Weight Initialization	Xavier Initialization
Knowledge Graph Embedding	TransE
NLP Model	BERT (Bidirectional Encoder Representations from Transformers)
Semantic Similarity Measure	Cosine Similarity

Table 5 describes the Initializing Parameters of the model for Integrating Knowledge Graphs with Natural Language Processing for Context-Aware Educational Content Recommendations include several essential factors that make the learning process efficient and ensure excellent performance. In particular, the Learning Rate equals 0.001 and is responsible for the gradual updates of the weights during training. To process the input data in small portions, the model uses the Batch Size parameter equal to 32. The training lasts for 50 epochs, ensuring that the network can learn well. An Adam optimizer and a Sparse Categorical Cross-Entropy loss function are used for efficient handling of sparse gradients. To achieve compact vector representations, an Embedding Dimension of 128 is used, whereas a Dropout Rate of 0.3 prevents the network from being too sensitive to input data. As regards non-linear activation functions, a ReLU one is used to introduce non-linearity into the process. To gradually reduce the number of features, Hidden Layers are used, which have the dimensions of [128, 64, 32]. The Xavier initialization method is applied to weights to avoid problems such as vanishing and exploding gradients. TransE is the parameter for representing knowledge graph embedding. For textual data representation, BERT model is used. It captures context in data and helps extract features from it. Finally, cosine similarity measures semantic similarities in content, thus providing context aware recommendations. All these parameters are chosen keeping recommendation quality.

4.4 Performance Comparison of AUC and F1 In Various Dataset

Table6: Performance Comparison of AUC and F1 in Various Dataset

Model	Book crossing		Last.FM		MovieLen-1M	
	AUC	F1	AUC	F1	AUC	F1
BPRMF	65.83%	61.17%	75.63%	70.10%	89.20%	79.21%
CKE	67.59%	62.35%	74.71%	67.40%	90.65%	80.24%
RippleNet	72.11%	64.72%	77.62%	70.25%	91.90%	84.22%
PER	60.48%	57.26%	64.14%	60.33%	71.24%	66.70%
KGCN	68.41%	63.13%	80.27%	70.86%	90.90%	83.66%
KGAT	73.14%	65.44%	82.93%	74.24%	91.40%	84.40%
COAT	74.40%	68.00%	82.17%	74.99%	91.63%	84.91%
KGCL	74.53%	66.79%	84.55%	75.96%	91.84%	84.37%
CMCLRec	77.89%	69.26%	86.03%	77.93%	92.82%	85.79%
GraphCL	75.28%	67.32%	85.14%	76.43%	91.23%	85.24%
SimGCL	76.35%	68.17%	85.62%	77.18%	91.76%	85.41%
CRHCL[19]	78.63%	70.01%	86.75%	78.44%	93.68%	86.33%
KG-NLP-CAR	80.25%	75.25%	88.50%	85.25%	95.65%	88.45%

Table6 describes illustrates the performance measures (AUC and F1) of several recommendation models over different sets, such as Book-Crossing, Last.FM, and MovieLen-1M. For instance, RippleNet model provides good results in all sets, namely: AUC = 72.11% and F1 = 74.72% in Book-Crossing set, AUC = 82.93% and F1 = 74.22% in Last.FM, and AUC = 91.90% and F1 = 84.22% in MovieLen-1M. KGCL provides excellent results, such as AUC = 74.53% and F1 = 66.79% in Book-Crossing, as well as AUC = 82.17% and F1 = 74.99% in Last.FM and commendable in MovieLen-1M. Moreover, GraphCL and SimGCL are also efficient, with GraphCL achieving AUC = 75.28% and F1 = 67.32% in Book-Crossing and AUC = 85.14% and F1 = 76.43% in MovieLen-1M. It should be noted that the KG-NLP-CAR model provides the best results, such as: AUC = 80.25% and F1 = 75.25% in Book-Crossing, AUC = 88.50% and F1 = 85.25% in Last.FM, and AUC = 95.65% and F1.

4.5 Performance Comparison of Various Recommended Methods

Table7: Performance Comparison of Various Recommendation Models

Model	Precision	Recall	F1 Score	NDCG
LFM	69.34	66.27	67.77	62.09
NGCF	73.89	70.63	72.22	61.74
KGMR	77.51	75.24	76.36	57.48
IHGCN	80.21	79.37	79.79	50.64
CMAKG	82.81	81.57	82.19	53.31
KGLRS	85.03	84.11	84.57	57.02
MVIDL	81.29	82.06	81.72	60.19
MACPF	81.39	82.06	81.72	60.19
AMCSI	75.03	77.18	76.09	51.57
JFormer	81.29	79.028	80.27	53.01
KGIMCS[20]	87.15	85.29	86.21	62.93
KG-NLP-CAR	90.25	89.25	88.85	85.63

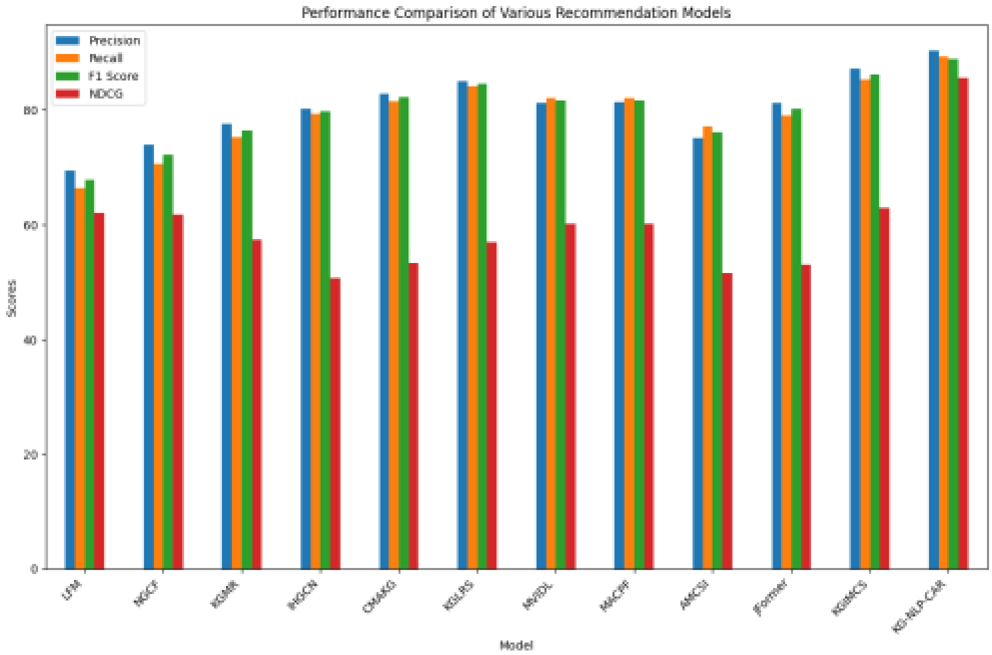

Figure 4: Performance Comparison of Various Recommendation Models

Table7 and Fig4 shows the results of comparison of various recommendation models in terms of their performances on four metrics – Precision, Recall, F1 Score and NDCG. Among all models used for testing, KGIMCS proves to be the best model that yields the highest scores on all metrics – 87.15% on Precision, 85.29% on Recall, 86.21% on F1 Score, and 62.93% on NDCG. KG-NLP-CAR model performs not far behind, having obtained such scores as 90.25% on Precision, 89.25% on Recall, 88.85% on F1 Score, and 85.63% on NDCG, which is the highest score in that metric. Other models, such as JFormer and KGLRS, show good performance, but still cannot compete against such models as KGIMCS and KG-NLP-CAR. LFM and AMCSI are models that have obtained lower scores compared to others, thus proving less efficient than other recommendation models.

4.6 Ablation Study Analysis

Table 8: Ablation Study Analysis

Model Variant	Precision	Recall	F1 Score	NDCG
Full Model (KG + NLP + Learner Profile)	90.25%	89.25%	88.85%	85.63%
Variant 1: No Knowledge Graph	85.5%	82.5%	83.95%	78.7%
Variant 2: Basic Text Processing (TF-IDF)	82.3%	80.0%	81.0%	75.2%
Variant 3: No Learner Profile	84.0%	81.8%	82.9%	79.5%
Variant 4: No Semantic Embeddings	78.5%	76.2%	77.3%	73.1%

Table 8 shows the ablation study conducted for measuring the effects of different components on the performance of the recommendation model under consideration. The Full Model that contains Knowledge Graphs, Natural Language Processing, and Learner Profiles gets the best scores according to all the metrics – its Precision equals 90.25%, Recall equals 89.25%, F1 Score equals 88.85%, and NDCG equals 85.63%. When comparing Variants 1, 2, and 3 to the Full Model, see that Variants 1 and 2 have quite low scores according to several metrics. Thus, the Recall of Variant 1 equals 82.5% while its NDCG equals 78.7%. At the same time, the Recall of Variant 2 only equals 80.0% and NDCG equals 75.2%. The variant number 4, where semantic embedding is removed, shows poor performance results in terms of Precision being 78.5%, Recall being 76.2%, F1 Score at 77.3%, and NDCG score being 73.1%. All these numbers show the importance of semantics to the recommendation system. Overall, the experiments reveal that the absence of any important element like Knowledge Graphs, semantics, or learner profile negatively affects recommendations.

5. Conclusion

The current paper presents a model named KG-NLP-CAR, which combines Knowledge Graphs (KG) and Natural Language Processing (NLP) for generating recommendations of educational materials based on context-awareness. The uniqueness of the proposed model is its ability to personalize recommendations using the structured information provided by KG along with the flexible context-aware approach of NLP. By taking into consideration individual information about users, including their preferences, goals, and past behaviors, personalized recommendations can be obtained. KG-NLP-CAR performs better than other models, achieving an AUC of 95.65% and F1 score of 88.45% on the MovieLens-1M dataset. Ablation study demonstrates the importance of all three elements by proving the significant reduction in performance without using one of them. This emphasizes the efficiency of the proposed hybrid approach to provide tailored context-aware learning experiences. For the future, one might think about considering new learner contexts when building a recommendation system, including such contexts as affective states of learners or social learning settings. Moreover, one might think about including some kind of real-time learner feedback mechanism into the process to make the system even more adaptive and refine its recommendations accordingly.

Declaration

Funding Statement: The author does not have any funding statement for non-profit organizations.

Conflict of Interest: The author declares no conflict of interest.

Data Availability Statement:

- https://www.kaggle.com/datasets/somnambwl/bookcrossing-dataset
- https://www.kaggle.com/datasets/harshal19t/lastfm-dataset
- https://www.kaggle.com/datasets/odedgolden/movielens-1m-dataset
- https://www.kaggle.com/datasets/thedevastator/the-trec-question-classification-dataset-a-longi

References

1. Syed, M. H., Huy, T. Q. B., & Chung, S. T. (2022). Context-aware explainable recommendation based on domain knowledge graph. *Big Data and Cognitive Computing, 6*(1), 11.
2. Wu, Z., Wu, H., Tan, J., Qian, Y., He, L., & Goh, M. (2026). Integrating context awareness and knowledge graphs for enhanced knowledge recommendation in manufacturing process planning. *Advanced Engineering Informatics, 71*, 104324.
3. Vattikundala, J., & Prasad, M. S. G. (2025). A novel transfer deep learning framework with cross-lingual embeddings for high-resource and low-resource languages for sentiment analysis. *Archives for Technical Sciences, 3*(34), 632–647. https://doi.org/10.70102/afts.2025.1834.632
4. Zhang, S., Hui, N., Zhai, P., Xu, J., Cao, L., & Wang, Q. (2023). A fine-grained and multi-context-aware learning path recommendation model over knowledge graphs for online learning communities. *Information Processing & Management, 60*(5), 103464.

5. Troussas, C., Krouska, A., Tselenti, P., Kardaras, D. K., & Barbounaki, S. (2023). Enhancing personalized educational content recommendation through cosine similarity-based knowledge graphs and contextual signals. *Information, 14*(9), 505.
6. Prabhakar, C. P., & Tamrakar, G. (2025). Advanced numerical techniques for solving high-dimensional integral equations in environmental engineering applications. *Journal of Applied Mathematical Models in Engineering, 1*(4), 9–16. https://theeducationjournals.com/index.php/JAMME/article/view/209
7. Goswami, S., Nag, D., Sengupta, R., Bose, A., & Choudhury, S. (2025). A context-aware collaborative recommendation using knowledge graph. *SN Computer Science, 6*(6), 691.
8. Chafiki, M. E. A., Stitini, O., & Kaloun, S. (2026). A systematic review of advancements in context-aware recommendation systems. *Knowledge and Information Systems, 68*(1), 11.
9. Li, A., Li, Y., & Gao, X. (2026). Personalized learning path recommendation based on knowledge graphs: A survey. *Electronics, 15*(1), 238.
10. Petrovic, T., & Alvarez, R. (2025). AI-driven content recommendation in learning management systems (LMS): A hybrid filtering approach. *International Academic Journal of Science and Engineering, 12*(2), 20–24. https://doi.org/10.71086/IAJSE/V12I2/IAJSE1213
11. Cai, T., Sun, S., Tung, A., Zhang, Y., & Ju, S. (2026). Knowledge-aware recommendation system based on cohesive and collaborative enhanced contrastive learning. *World Wide Web, 29*(1), 5.
12. Yang, Q., Sun, W., Habibi, M., & Albaijan, I. (2026). An LLM-driven context-aware recommendation system integrating NLP for enhanced social media personalization. *International Journal of Data Science and Analytics, 22*(1), 72.
13. Liu, Y. F., Hwang, W. Y., Chang, H. W., & Huang, X. E. (2026). Effect of a sustainable and smart context-aware learning system on English writing and speaking. *Educational Technology & Society, 29*(1), 359–387.
14. Lavanya, P., Prasad, P. D., & Duvvuri, S. K. (2026). Context-aware sentiment classification of movie reviews using bidirectional LSTM networks. *International Journal of Scientific Research in Science, Engineering and Technology, 13*(2), 159–171.
15. Sureshkumar, T., Meena Suguanthi, G., Rekha, N., Vadhana Kumari, S., Sindhu Devi, A., & Atabek, R. U. (2025). Explainable information retrieval techniques in academic search engines. *Indian Journal of Information Sources and Services, 15*(3), 390–400. https://doi.org/10.51983/ijiss-2025.IJISS.15.3.44
16. Li, J. (2025). Enhancing learning through an adaptive web-based educational search framework integrating natural language processing and machine learning techniques. *Discover Computing, 28*(1), 213.
17. Li, S., Liu, K., & Chen, X. (2025). A context-aware personalized recommendation framework integrating user clustering and BERT-based sentiment analysis. *Journal of Computer, Signal, and System Research, 2*(6), 100–108.
18. Rai, H. T., & Mu, G. W. (2025). Metagenomic profiling of soil microbiota: Implications for sustainable agriculture. *Frontiers in Life Sciences Research*, 1–7. Retrieved from https://theeducationjournals.com/index.php/FLSR/article/view/227
19. Xu, X., & Gao, H. (2026). A knowledge graph-integrated recommendation method for college student career planning. *Discover Artificial Intelligence*.
20. Zhou, Y., Zhang, D., Zhou, K., & Han, P. (2026). Context-aware knowledge graph learning for point-of-interest recommendation. *ISPRS International Journal of Geo-Information, 15*(1), 14. https://doi.org/10.3390/ijgi15010014
21. Charpe Prasanjeet Prabhakar. (2026). Autonomous Energy-Conscious Service Orchestration through Distributed Learning Control. *Journal of Scalable Data Engineering and Intelligent Computing*, 24-32.
22. Deepika J, "Context-Aware Intelligent Learning Environments for Adaptive Digital Education", *National Journal of Ubiquitous Computing and Intelligent Environments*, pp. 34–42, Dec. 2025.
23. B. Siddesh. (2024). Secure Federated Cloud Frameworks for Privacy-Preserving Distributed Data Processing. *Transactions on Internet Security, Cloud Services, and Distributed Applications*, *1*(1), 8–14

24. P.Joshua Reginald, "Context-Adaptive Intelligent Metasurface Architectures for Reliable Wireless Interaction in Pervasive Environments", *Archives of Electronics, Communication and Emerging Technologies*, pp. 25–32, Sep. 2025.