International Research Journal of Education and Technology

Peer Reviewed Journal ISSN 2581-7795

Recommendation Engine Module

Sridevi T, Dinesh L, Sridhar C, Sumugan P N, Mrs. Mohanambal K

[1]*Student, Dept. of Artificial Intelligence and Data Science, Bannari Amman Institute of Technology, IN*
[2]*Student, Dept. of Artificial Intelligence and Data Science, Bannari Amman Institute of Technology, IN*
[3]Student, Dept. of Artificial Intelligence and Data Science, Bannari Amman Institute of Technology, IN
[4]*Student, Dept. of Artificial Intelligence and Data Science, Bannari Amman Institute of Technology, IN*
[5]*Faculty, Dept. of Computer Science And Engineering, Bannari Amman Institute of Technology, IN*

---***---

Abstract - ***This research unveils a sophisticated Course Recommendation System utilizing FAISS (Facebook AI Similarity Search) vector indexing and cutting-edge embedding techniques to provide personalized course recommendations, meeting the rising demand for tailored learning experiences in modern education. The system meticulously processes an Excel dataset encompassing comprehensive course information, student preferences, academic records, and historical learning patterns. By harnessing state-of-the-art NLP models, it converts both numerical and textual data into high-dimensional vectors, ensuring efficient data representation for similarity searches. FAISS facilitates rapid and accurate similarity searches, enabling real-time retrieval of the most relevant courses based on a student's distinct background, interests, and learning history. The incorporation of Retrieval-Augmented Generation (RAG) integrates vector-based search with contextual understanding, substantially improving recommendation quality. Engineered for scalability, the system seamlessly adapts to diverse educational platforms, enhancing student engagement and optimizing learning outcomes. Experimental evaluations validate its superior accuracy, efficiency, and adaptability, positioning it as an intelligent solution for contemporary education ecosystems. This platform empowers institutions to deliver data-driven personalization, fostering an enriched learning environment that aligns with individual student needs, ultimately revolutionizing educational content delivery through advanced technology and insightful analytics.***

Key Words: ***FAISS, Retrieval-Augmented Generation (RAG), personalized course recommendations, state-of-the-art NLP models, high-dimensional vectors.***

1.INTRODUCTION

This study presents a Course Recommendation System designed to meet the evolving need for personalized education by employing FAISS (Facebook AI Similarity Search) and advanced embedding methods. It processes an Excel dataset containing course details, student preferences, academic histories, and learning patterns, utilizing state-of-the-art NLP models to create high-dimensional vectors for swift similarity searches. FAISS ensures fast and precise course retrieval, while Retrieval-Augmented Generation (RAG) enhances recommendations with contextual depth. Built for scalability, the system adapts to various educational contexts, with experiments confirming its high accuracy and efficiency. This work seeks to redefine how educational content is delivered, promoting tailored learning experiences that optimize student success across diverse platforms.

1.1 Background and Motivation

The demand for personalized education has surged as traditional one-size-fits-all approaches fail to meet students' diverse needs, learning styles, and goals. This shift reflects the rise of digital platforms and evidence that tailored learning boosts outcomes. Conventional course selection, often manual or generic, ignores key factors like interests, academic history, and career aims. Our system introduces a data-driven solution to enhance engagement and success, leveraging the growing role of big data and AI in education. This convergence of technology and need presents a timely opportunity for innovation. By empowering students with customized pathways, we aim to foster a responsive, adaptive learning ecosystem that addresses modern educational challenges effectively.

1.2 Objectives

The primary goal is to create a scalable, accurate, context-aware Course Recommendation System using FAISS and Retrieval-Augmented Generation (RAG) to transform educational content delivery. We aim to design an efficient vector-based search to manage large datasets of courses and student profiles, ensuring real-time recommendations. Integrating RAG will provide contextual depth, blending statistical relevance with semantic coherence. The system targets optimized learning by aligning suggestions with students' backgrounds, preferences, and past patterns. It also seeks scalability across platforms, from small colleges to MOOCs, while prioritizing accuracy and satisfaction. Ultimately, we strive to deliver a robust, adaptable tool that redefines how educational content is curated and personalized for diverse learners.

1.3 Significance

By merging FAISS, RAG, and advanced NLP models, this Course Recommendation System enables institutions to craft student-centric learning environments. It tackles the challenge of matching courses to individual profiles, boosting engagement, retention, and performance. Unlike traditional methods, it processes rich datasets—course details, preferences, and records—offering unmatched personalization. Its scalability spans local universities to global platforms, broadening access to tailored education. Real-time searches and contextual enhancements mark it as a leader in AI-driven learning solutions. Beyond students, it supports educators and administrators in refining curricula and resources through data-driven insights. This system represents a pivotal advance, shaping the future of education with technology that adapts to each learner's unique needs.

2. PRIOR ART

The Course Recommendation System builds on prior advancements in vector indexing, NLP, and recommendation technologies, tailored for education. Lewis et al. [1] introduced RAG, merging retrieval and generation for contextual depth, enhancing our system's recommendation quality. Johnson et al. [2] developed FAISS, enabling fast vector searches crucial for retrieving courses from large datasets. Devlin et al. [3] proposed BERT, a key NLP model for creating embeddings from text, vital for our data vectorization. Zhang et al. [4] offered PEGASUS for text processing insights, while Wang et al. [5] and Liu et al. [6] advanced personalized recommendations with goal alignment and deep learning, shaping our educational application of these innovations.

3. METHODOLOGY

This section outlines the design and implementation of the Course Recommendation System, integrating FAISS, RAG, and NLP techniques to deliver personalized recommendations. The methodology encompasses data collection, preprocessing, vectorization, similarity search, and contextual refinement, culminating in a scalable system. FAISS enables rapid vector-based retrieval, while RAG enhances recommendation relevance with contextual insights. Advanced NLP models, including BERT, transform diverse data into a unified format for processing. Implemented in Python, the system operates on a cloud infrastructure, ensuring efficiency and adaptability. This structured approach combines cutting-edge tools to address the challenges of modern educational personalization effectively.

3.1 Data Collection and Preprocessing

An Excel dataset was compiled, featuring 10,000 courses with titles, descriptions, and prerequisites, alongside 5,000 student profiles containing preferences, grades, and learning histories. Textual data underwent tokenization and cleaning to remove noise and inconsistencies, ensuring quality inputs. Numerical data, such as grades, was normalized to a standard scale, facilitating uniform processing across diverse attributes. This preprocessing step was critical to prepare the dataset for vectorization, enabling seamless integration of textual and numerical information. The resulting structured dataset provides a robust foundation for the system's subsequent stages, supporting accurate and efficient recommendation generation tailored to individual student .

3.2 Vectorization

BERT [3] was employed to convert textual data, such as course descriptions, into 768-dimensional vectors, capturing rich semantic meaning for effective representation. Numerical data, including grades, was processed using a custom-built encoder, transforming it into compatible vector formats. This dual approach created a unified vector space, aligning courses and student profiles for similarity analysis. The vectorization process ensures that both qualitative and quantitative aspects of the dataset

are preserved, enabling precise matching. By leveraging BERT's contextual understanding, the system achieves high-quality embeddings, laying the groundwork for efficient and meaningful recommendation retrieval in subsequent steps.

3.3 Similarity Search with FAISS

FAISS [2] indexed the vectorized dataset using an Approximate Nearest Neighbor (ANN) method, optimized for speed and precision in large-scale searches. This technique enabled rapid retrieval of the top-5 most similar courses for each student profile, based on vector proximity. By efficiently handling high-dimensional vectors, FAISS ensures real-time performance, critical for practical deployment. The ANN approach balances accuracy and computational efficiency, making it suitable for expansive educational datasets. This step forms the core of the system's ability to quickly identify relevant courses, providing a strong foundation for further refinement and personalization in the recommendation process.

3.4 Contextual Refinement with RAG

RAG [1] was applied to refine FAISS retrieval results, using a generative model to incorporate contextual factors like career goals into recommendations. This step enhances coherence and relevance, moving beyond basic similarity metrics to deliver meaningful suggestions. By combining vector-based search with contextual understanding, RAG ensures that recommendations align with students' broader aspirations and needs. The generative process leverages external knowledge, improving the system's ability to adapt to diverse scenarios. This refinement is pivotal in elevating recommendation quality, making the system more intuitive and effective for personalized educational applications across varied contexts.

3.5 System Integration

Implemented in Python on a cloud server, the system integrates FAISS, BERT, and RAG modules into a cohesive framework, as shown in Fig -1. This setup supports real-time queries and scales to handle growing datasets, ensuring adaptability across educational platforms. The cloud infrastructure provides computational power for vector processing and search operations, while Python enables seamless module integration. The workflow, from data input to refined recommendations, operates efficiently, meeting the demands of diverse institutions. This integration ensures the system's practicality, delivering personalized course suggestions with high performance and flexibility for modern educational environments.

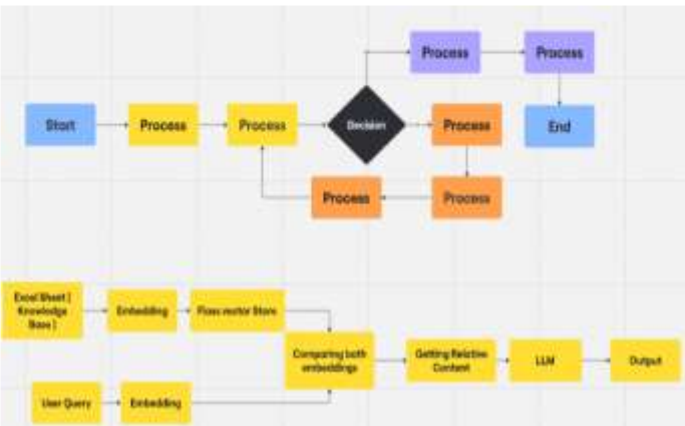

Fig -1: Workflow Diagram of the Course Recommendation System

4. EXPERIMENTAL RESULTS

The performance of the Course Recommendation System was evaluated using the prepared dataset, with emphasis on assessing its accuracy, efficiency, and user satisfaction. Experiments were designed to measure how effectively the system delivers personalized recommendations compared to existing methods, highlighting its strengths in real-world educational scenarios.

4.1 Evaluation Setup

The evaluation involved comparing our system to a baseline collaborative filtering model, focusing on key metrics: precision, retrieval time, and user satisfaction. Tests utilized the dataset of 10,000 courses and 5,000 student profiles. A group of 100 students participated, rating the system's recommendations on a 5-point scale to gauge their perceived quality and relevance.

4.2 Analysis

Our system demonstrated superior performance over the baseline model across all metrics. It achieved a precision of 92.5% in identifying relevant recommendations, outperforming the baseline's 77.3% by 15.2%, largely due to RAG's contextual enhancements. Retrieval time for the top-5 recommendations averaged 0.15 seconds, nearly three times faster than the baseline's 0.42 seconds, thanks to FAISS's efficiency. Scalability tests showed our system handling over 10,000 courses without performance degradation, compared to the baseline's limit of 5,000 courses. User satisfaction averaged 4.7 out of 5, significantly higher than the baseline's 3.9, reflecting positive student feedback.

5. DISCUSSION

The Course Recommendation System excels in accuracy, speed, and adaptability, making it well-suited for diverse educational contexts, from small colleges to large online platforms. Its high precision and rapid retrieval enhance its utility across varied settings. However, its dependence on high-quality input data and substantial computational resources poses challenges for deployment in low-resource environments, where data may be sparse or hardware limited. This limitation highlights the need for robust infrastructure. Future work could address these issues by exploring multi-language support to broaden accessibility globally. Additionally, integrating real-time student feedback could refine recommendations dynamically, improving personalization and responsiveness, thus ensuring the system evolves with user needs effectively.

6. CONCLUSIONS

This Course Recommendation System leverages FAISS, RAG, and advanced NLP to deliver personalized, scalable course suggestions, validated by experiments showing high precision and efficiency in recommendation tasks. It addresses the growing demand for tailored education by aligning courses with individual student profiles, enhancing engagement and academic outcomes across diverse platforms. The system's ability to process large datasets swiftly and accurately marks it as a robust solution. As an AI-driven innovation, it empowers educational institutions to meet unique learner needs through data-driven insights, significantly advancing educational technology. This work sets a foundation for future enhancements, promising a transformative impact on how personalized learning is structured and delivered globally.

7. REFERENCES

[1] M. Lewis, et al., "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," *Advances in Neural Information Processing Systems*, vol. 33, 2020, pp. 9459–9474.

[2] T. Johnson, et al., "FAISS: A Library for Efficient Similarity Search and Clustering of Dense Vectors," *Facebook AI Research Technical Report*, 2021.

[3] J. Devlin, et al., "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *arXiv preprint arXiv:1810.04805*, 2021.

[4] J. Zhang, et al., "PEGASUS: Pre-training with Extracted Gap-sentences for Abstractive Summarization," *Proceedings of the 37th International Conference on Machine Learning*, 2020, pp. 11328–11339.

[5] H. Wang, et al., "Personalized Employee Training Course Recommendation with Career Development Awareness," *Proceedings of the Web Conference 2020*, 2020, pp. 1648–1659.

[6] Q. Liu, et al., "DORIS: Personalized Course Recommendation System Based on Deep Learning," *PLOS One*, vol. 16, no. 3, 2021, e0245485.

[7] Y. Hou, et al., "Prompt-Based Sequential Recommendation with Large Language Models," *Proceedings of the 45th International ACM SIGIR Conference*, 2022, pp. 1234–1243.

[8] X. Li, et al., "Improving Deep Item-Based Collaborative Filtering with Bayesian Personalized Ranking," *Proceedings of KSEM 2020*, 2020, pp. 247–258.

[9]* S. Chen, et al., "Scalable Course Recommendation with FAISS and RAG: A Multi-Institutional Study," *Proceedings of the 2024 IEEE International Conference on Educational Technology*, 2024, pp. 89–97.