Grenze International Journal of Engineering and Technology, June Issue

A Hybrid Framework Integrating LLM and RAG for Personalized and Adaptive Skill Roadmap Generation in Career and Learning Pathways

Sumathi D[1] and M D Ganesha[2]
[1-2]Department of Computer Science and Engineering, CMR Institute of Technology, Bangalore, India.
Email: d.sumathiphd@gmail.com

***Abstract*— As technology continues to develop rapidly due to the availability of information technology, learners are in need of developing their skills. Hence, learners frequently encounter unstructured material and finding simple routes through which they can learn new skills. This research provides a framework that integrates AI-enabled tools along with large language models (LLMs) into a retrieval-augmented generation (RAG) framework, for enabling the generation of personalized skill development maps to provide required information and materials. By using these methods, users can pursue learning pathways across multiple domains, based on each user's current level of skill, their individual goal, with their available time. With the locally hosted Mistral-7B-Instruct model with the Ollama framework, this study prioritizes privacy, efficiency, and the ability for users to create their learning pathways offline. Contextual retrieval capabilities are provided through a Pinecone vector database that uses 384-dimensional sentence transformer embeddings built with all-MiniLM-L6-v2, which supports a multi-phase pipeline that includes dataset curation, prompt preparation, retrieval, and the generation of an interactive roadmap. Through this research, we contribute to a privacy preserving solution with salability. It will helpful for adaptive career development and also the institutional training programs. It reduces the time of learning track discovery to 65% compared to manual approach. The personalization quality and experimental evaluation accuracy is 80%, precision is 82%and measured recall is 79%. The overall average user satisfaction score has increased to 4.2 out of 5.0.**

***Index Terms*— Adaptive Skill Roadmap, Hybrid LLM-RAG Framework, Retrieval-Augmented Generation, Intelligent Recommendation Systems, Dynamic Curriculum Design, Career Development.**

I. INTRODUCTION

Technological development has increased day by day. The rapid changes in technology and revolution in tools advancements are increased. Learning and/or updating advanced technologies is a challenging one for student learners and working professionals, also employers. Identifying the individual's learning requirements based on their interest and their's efficiency is challenging one. The traditional methods are very general, which gives all sources, providing unnecessary materials. This will make an individual in confusion stare and difficult to identify the proper path based on their interest and efficiency. The recent advancement in Large Language Maodels (LLMs) have a fantastic capability in natural language understanding.

Grenze ID: 01.GIJET.12.2.140_23

The LLMs useful for multi-lingual capabilities, contextual reasoning and knowledge synthesis [7]. However, existing LLM-based career recommendation systems exhibit several limitations: tendency toward hallucination and generation of factually incorrect information when operating without grounded knowledge sources; inability to access domain-specific, up-to-date information beyond training data cutoffs; lack of personalization mechanisms that account for individual learner profiles; and insufficient integration with structured knowledge bases [2,3,4,6].

Experts Rule-Based Systems will provide a structure (Recommendation); however, they will not be flexible for different people; they require a lot of manual maintenance and will not change automatically based on new skill sets or industry requirements [12]. Furthermore, online learning platforms are generic; therefore, thousands of options would overwhelm a user with no personalization, formal learning pathway for a more efficient learning experience.

The proposed study addresses these drawbacks by suggesting a combination of the advantages associated with the generation of content using LLMs and the information supported by RAG. In addition, this research uses a curated, subject-specific (domain- specific) set of structures to develop and create contextual and factual learning road maps for users.

II. Literature Survey

This survey of existing systems presents a review in AI-driven personalized learning and skill development systems suitable for individuals, and also analyze the career guidance platforms. The survey identifies research gaps, evaluates technological approaches based on the educational systems, and contextualizes the proposed framework in educational technology.

A. AI in Educational Systems

Jamal et al. [1] AI's role in teacher education was re- searched, the capacity of implementing an adaptive instructional model combined with providing immediate feedback for both students and instructors (or mentors) through adaptive and personalized skill development using additional methods. Martin and Sumithra [2] A generative AI system was developed for students for job placement by creating personalized career maps in real time. Using natural language processing to analyze student profiles, the generative AI generated a personalized learning path for each student and resolved the gaps between their preparedness. Razaque et al. [3] A study was conducted to explore how Artificial Intelligence can improve the design of user interfaces for online learning environments. The specific areas of research included Adaptive Navigation, Gamification, and Predictive Learning Path Recommendations.

B. Large Language Models in Education

Numerous systematic reviews have been completed that discussed both ChatGPT and similar GenAI[4,5] advantages and disadvantages regarding education. These reviews generally highlight the benefits of using AI as personal tutors, generating materials for course design, and automating administrative tasks; nevertheless, each of the reviews expressed worries about the cultural sensitivity of the technology, potential biases in generated content, factual correctness, and alignment with existing curricula.

Memarian and Doleck [6] through a systematic review of 63 studies examining ChatGPT's effects in terms of education, the authors concluded that although LLMs do provide opportunities for personalisation and improved efficiency in education, they also create a range of issues regarding Academic Integrity, Data Privacy, Transparency and over-reliance upon AI-generated content.

Qadir [7] The research analysis on engineering education on Generative AI found potential advantages in three areas of using the technology: Content Creation, Tutoring and Assessment, and outlining potential disadvantages including issues with Plagiarism or Misrepresentation or any other types of Academic Dishonesty; Superficial Learning; Factual Inaccuracies.

C. Personalized Learning Systems

Garc´ıa-Pen˜alvo [8] The research demonstrates that GenAI has multiple applications in education. The ability to automate content creation, personalize learning experiences, improve administrative efficiency and identify challenges associated with sustainability. These include gaps in AI literacy, ethical concerns, environmental impact and data privacy issues associated with the use of AI technology in education.

Murtaza et al. [9] The study examined how machine learning, deep learning, and knowledge tracing (KT) are used by AI-Based Personalized e-Learning Systems (AI-PLS). The framework does improve engagement, personalization, and the quality of the learner profile, however, it is limited by content diversity, generalizability across different domains and difficulty scaling for large-scale implementation. A comprehensive review by D.G.M et al. [10] A review of the study of personalized learning recommendation systems for online education based on 60 articles demonstrates increased learner engagement and retention through adaptive technologies like Fluxy AI and chatbots.

D. AI-Powered Mentorship and Career Guidance

The author Bagai and Mane [11] explains in his research and explained how the MentorAI works as tools for professional development of individual. This study explained that machine learning and natural language processing. Hence it became a powerful and helps to deliver useful and real-time support. But the difficulty is, it raised some important apprehensions like bias in algorithms. The struggle to truly capture human emotions with code, and then the worries about data privacy. The authors Bhatt and Muduli [12] did a review of 81 papers on AI in Learning and Development. It explains a real progress in, how do we design the system, training, testing and validating then deliver it.

E. AI Tutoring and Assessment Systems

In another research done by Frankford et al. [13], they developed an AI-Tutor concept using the large language model and GPT-3.5 Turbo. This is very much useful in teaching software engineers for developing their programming skills. Testing knowledge of learners got more involved in programming logics. The learners themselves can managed their own learning. The learners then immediately got feedback so that the learners can analyze and improve their learning efficiency. Chukwuere [14] analyzed about the efficiency of ChatGPT in higher education. But the ChatGPT failed on human emotional cues. The ChatGPT reduces the analytic reasoning skill of learners.

III. PROPOSED METHODOLOGY

The proposed methodology presents a comprehensive description of the proposed hybrid framework for personalized skill roadmap generation. This section includes the system architecture, data processing pipeline, retrieval-augmented generation mechanism, personalization algorithms, and user interface design.

A. System Architecture

The multi-tier architecture consisting of four primary components: (1) Data Layer, (2) Processing and Retrieval Layer, (3) Generation and Personalization Layer, and (4) Presentation Layer. Figure 1 illustrates the complete system architecture. The Data Layer is a prepared library of structured materials for learning about 4 subjects: Web Development; Data Science; Cybersecurity DevOps. For each resource entry, the following information is provided: resource difficulty (beginner, intermediate, advanced), estimated time necessary for study, any prerequisites that must be completed prior to study, the expected learning outcomes, and citation(s) for the resource.

B. Dataset Curation and Preprocessing

Learning resources from officially recognized educational sources were compiled systematically to create the knowledge base. The knowledge base consists of 1247 individual learning modules divided between the four target domains. Resources were collected from the sources like, official technology documentation, recognized online learning platforms, professional certification curricula, open-source learning repositories.

C. Local LLM Deployment

The system uses Mistral-7B-Instruct, which is deployed locally via the Ollama framework. It ensures data privacy and offline capability and no need of third-party cloud APIs. It constrained prompts and enforce phase-wise sequencing, prerequisite awareness, and also time allocation.
Let $D = \{d_1, d_2, \ldots, d_N\}$ denote the curated knowledge base containing $N = 1247$ learning modules. Each module $d_i$consists of textual content and structured metadata (domain, difficulty, prerequisites, learning outcomes, and estimated time). After preprocessing, each module is transformed into a dense semantic embedding using a sentence transformer encoder:

$$e_i = f_{enc}(d_i; \theta_{enc}), e_i \in R^{384}$$

where f_{enc} represents the all-MiniLM-L6-v2 encoder and θ_{enc} denotes its parameters. All embeddings are indexed in a Pinecone vector database for efficient similarity search. A user query q, composed of current skill level, learning goals, and time constraints, is encoded using the same embedding function:

$$e_q = f_{enc}(q; \theta_{enc})$$

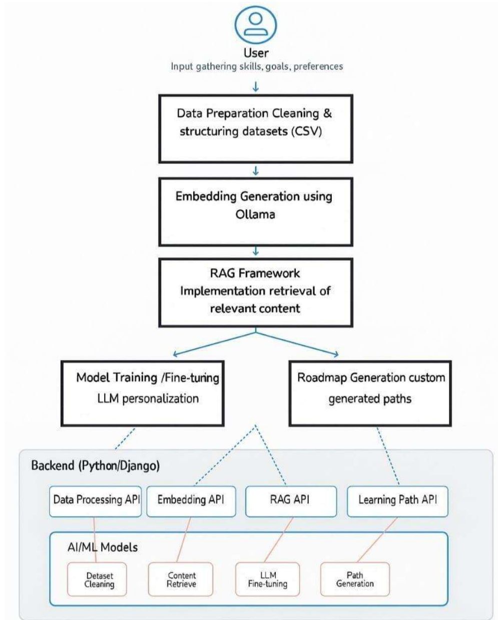

Fig. 1. System architecture

Similarity between the user query and each knowledge base entry is computed using cosine similarity:

$$sim(e_q, e_i) = \frac{e_q \cdot e_i}{\parallel e_q \parallel \parallel e_i \parallel}$$

The retriever selects the top-k most relevant modules: $R = top - k\ \{sim(e_q, e_i)\}, k \in [5,15]$

Logical namespace partitioning based on domain and difficulty level further improves retrieval accuracy and reduces latency.

The prompt template structure is designed as per the following figure 2.

```
   You are an expert career advisor. Based on:
- User Profile: {skill_level, domain, goals}
- Available Time: {time_constraint}
- Retrieved Resources: {resource_summaries}

   Generate a structured learning roadmap with:
1.Clear learning phases
2.Specific skills to acquire in each phase
3.Recommended resources
4.Time estimates
5.Prerequisites and dependencies

   Format output as structured JSON.
```

Fig 2: Prompt Template Structure

D. Personalization and Sequencing Logic

The personalization and sequencing logic is implemented based learner profile. The model incorporating from skill level of learner, learning goals, available learning time. It also analyzes the prerequisite knowledge of the learner. The retrieved modules suitable for the learners are first filtered and then ranked based on suitability. This ranking uses a weighted score reflects three aspects, the difficulty suitability, prerequisite satisfaction, and goal relevance. This ensures that the sequencing of the modules is logically correct. A time-constrained compression mechanism eliminates the low-priority modules when the learning time exceeds the availability of the user. This ensures that the roadmaps are feasible and learner specific.

All the learners are demonstrated as a tuple:

$$u = (s, g, t, p)$$

Were,

- s denotes the current skill level.
- g is denoting the learning goal.
- t is denoting the available learning time. and
- p is denoting the set of known prerequisites.

The levels of skill are encoded numerically as for beginner=0, intermediate=1, and advanced=3. The skill gap is computed as:

$$\Delta L = L_{target} - L_{current}$$

This gap regulates the difficulty path of students. This gap in learning level is useful for developing the roadmap.

E. Backend Integration and API Design

FastAPI is the backend of the system. The FastAPI is used to retrieve the information, LLM inference and personalization. The user profiles and feedback are stored in SQLite. The knowledge base is kept in a structured JSON format to make modifications in future and portability easier. The RESTful APIs are used to handle user input, producing the roadmap and collecting response and criticism. The front end is done via Streamlit.

Module Ranking and Selection

Each retrieved module $m_i \in R$ is scored using a weighted alignment function:

$$\alpha(u, m_i) = w_1\alpha_{diff} + w_2\alpha_{prereq} + w_3\alpha_{goal}$$

where α_{diff}measures difficulty suitability, α_{prereq}evaluates prerequisite satisfaction, and α_{goal}assesses relevance to user goals. The weights $(w_1, w_2, w_3) = (0.4,0.3,0.3)$were empirically chosen. Modules are selected if:

$$\alpha(u, m_i) \geq \theta, \theta = 0.6$$

The system was evaluated on standard commodity hardware (Intel Core i7, 16 GB RAM), achieving an average roadmap generation latency of 4.2 seconds, demonstrating suitability for real-time interactive use.

F. System Integration and Performance

This entire process is orchestrated by a FastAPI backend, which handles the retrieval, personalization, and LLM inference process. Additionally, the user profiles and feedback are stored using SQLite, and the entire frontend has been built using Streamlit. Finally, the entire system has been deployed on commodity hardware, running an Intel Core i7 processor, 16 GB of RAM, and has an average latency of 4.2 seconds for generating the roadmap, validating the entire system for a real-world deployment.

Fig. 3. explains the system architecture diagram illustrating a complete deployment pipeline and technology stack. As depicted by the diagram, the entire system has four main components: BASE MODEL and MODEL layer, which contains the OLLAMA base model, the Mistral-7B, and the API for the OCR process, the Training Dataset layer, which uses the web scraping and synthetic data generation for the fine-tuning process, the VECTOR Database layer, which uses the Chroma DB for the local development and the Pine Cone for the actual production, and the HUGGING FACE, which handles the GGUF file storage, the deployment, and the storage of the model, connected to the Production Endpoint and the Backend for the user data storage.

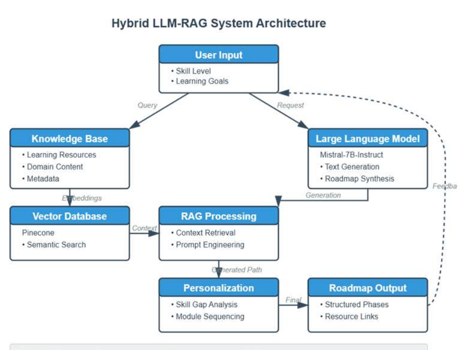

Fig. 3. The LLM and RAG integrated structure illustrating a complete deployment pipeline and technology stack

Fig 4 illustrates the explanation on the Sample FastAPI JSON Response on the generated structure on the roadmap Response includes: (a) metadata with user profile parameterization and timestamp on generation, (b) structured learning phases with particular skills and resources used, (c) time allocations for each module, and (d) prerequisites. The JSON format enables a seamless integration with frontend visualization components and supports export functionality for external learning management systems.

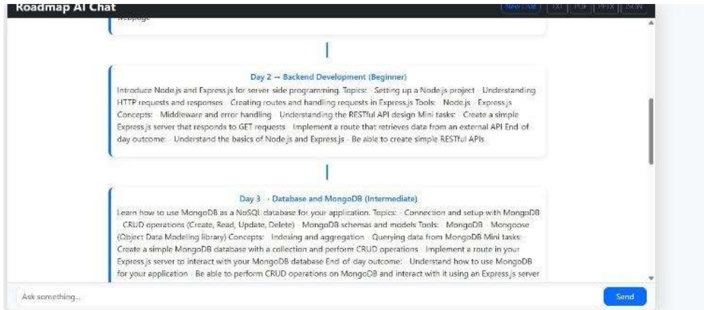

Fig. 4. The Roadmap-AI-Chat displays a Roadmap to follow for the learner

Table I presents performance breakdown across the four target domains. Overall, the results from each of the areas demonstrate a high degree of consistency across all of the different disciplines.

TABLE I: PERFORMANCE METRICS BY DOMAIN

Metric	Web Dev	Data Sci	Cyber	DevOps
Accuracy	82.1%	81.3%	78.2%	79.8%
Precision	84.3%	82.7%	79.5%	81.2%
Recall	80.5%	81.1%	76.8%	78.6%
F1-Score	82.3%	81.9%	78.1%	79.8%

III. CONCLUSIONS

This research article proposed a hybrid framework that integrates LLM and RAG to produce a customized adapted skill roadmaps for a person's career and learning pathways to upgrade these skills. By providing a generative outputs suitable for a learner, a domain-specific knowledge base and enabling local deployment. Hence, the system improves factual accuracy, personalization, and data privacy. The experimental results demonstrates a strong performance, about 80% roadmap accuracy with improved user satisfaction. It also reduces the response latency. This research produces the effective retrieval-grounded generative models for reliable AI-driven learning guidance based on the persons' pervious knowledge. The system also highlights their potential for scalable deployment in educational and professional training environments.

REFERENCES

[1] A. Jamal, "The role of artificial intelligence (AI) in teacher education: Opportunities & challenges," International Journal of Research and Analytical Reviews, vol. 10, no. 1, pp. 386–394, 2023.

[2] S. Martin and A. Sumithra, "Personalized AI-driven roadmap genera- tion," International Journal for Research Trends and Innovation, vol. 10, no. 1, pp. 120–128, Jan. 2025.

[3] A. Razaque, S. Hariri, and J. Yoo, "AI-driven user interface design: Enhancing digital learning and skill development," International Journal of Human-Computer Studies, vol. 186, article 103195, 2025.

[4] K. Achour, M. D. Laanoui, and M. Ourahay, "The impact of ChatGPT in education: A comprehensive overview," in Proc. International Conference on Global Aeronautical Engineering and Satellite Technology (GAST), Marrakech, Morocco, 2024, pp. 1–6.

[5] A. A. Fenta, "A review on enhancing education with AI: Exploring the potential of ChatGPT, Bard, and generative AI," Discover Education, vol. 4, article 38, 2025.

[6] B. Memarian and T. Doleck, "ChatGPT in education: Methods, potentials, and limitations," Computers in Human Behavior: Artificial Humans, vol. 1, no. 2, article 100022, 2023.

[7] J. Qadir, "Engineering education in the era of ChatGPT: Promise and pitfalls of generative AI for education," IEEE Transactions on Education, vol. 67, no. 3, pp. 250–265, Aug. 2024.

[8] F. J. Garc´ıa-Pen˜alvo, "Generative artificial intelligence and education: An analysis from multiple perspectives," Education and Knowledge Society Journal, vol. 25, article e31867, 2024.

[9] M. Murtaza, Y. Ahmed, J. A. Shamsi, F. Sherwani, and M. Usman, "AI-based personalized e-learning systems: Issues, challenges, and so- lutions," IEEE Access, vol. 10, pp. 81323–81339, 2022.
[10] D. G. M., R. H. Goudar, A. A. Kulkarni, V. N. Rathod, and G. S. Hukkeri, "A digital recommendation system for personalized learning to enhance online education: A review," in Proc. IEEE International Conference on Intelligent Systems for Communication, IoT and Security (ICISCoIS), Coimbatore, India, 2023, pp. 245–250.
[11] R. Bagai and V. Mane, "Designing an AI-powered mentorship platform for professional development: Opportunities and challenges," Interna- tional Journal of Computer Trends and Technology, vol. 71, no. 4, pp. 108–114, Apr. 2023.
[12] E. Frankford, C. Sauerwein, P. Bassner, S. Krusche, and R. Breu, "AI- tutoring in software engineering education: Experiences with large lan- guage models in programming assessments," in Proc. ACM Conference on Innovation and Technology in Computer Science Education (ITiCSE), Milan, Italy, 2024, pp. 412–418.
[13] J. E. Chukwuere, "The use of ChatGPT in higher education: The ad- vantages and disadvantages," Journal of Applied Learning and Teaching, vol. 6, no. 1, pp. 24–31, 2023
[14] P. Bhatt and A. Muduli, "Artificial intelligence in learning and development: A systematic literature review," European Journal of Training and Development, vol. 47, no. 7/8, pp. 677–694, 2023.