



International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)





GenAI-Based Career Guidance Tool for Rural Students

Dr. Meghana G R, Preethu B T, Priyanka R, Prakruthi N, Sneha S Ballary

Dept. of Computer Science & Engineering, Jain Institute of Technology, Davangere, India

Dept. of Information Science & Engineering, Jain Institute of Technology, Davangere, India

Dept. of Information Science & Engineering, Jain Institute of Technology, Davangere, India

Dept. of Information Science & Engineering, Jain Institute of Technology, Davangere, India

Dept. of Information Science & Engineering, Jain Institute of Technology, Davangere, India

ABSTRACT: Career guidance remains a critical yet significantly underserved domain for students in rural and semi-urban regions of developing nations. Traditional counseling approaches suffer from geographical constraints, language barriers, and a severe shortage of qualified career advisors. This paper presents a novel Generative Artificial Intelligence (GenAI)-powered career guidance system tailored specifically for rural students, leveraging Large Language Models (LLMs), Natural Language Processing (NLP), and multi-modal interaction to deliver personalized, context-aware career recommendations. The proposed system, termed RuralCareerGPT, integrates a fine-tuned transformer-based model with a Retrieval-Augmented Generation (RAG) pipeline, enabling it to respond intelligently to student queries in regional languages. Evaluated on a curated dataset of 4,200 rural student profiles across five Indian states, RuralCareerGPT achieves a career recommendation accuracy of 91.4%, outperforming existing baseline systems by a margin of 7–14%. The system also demonstrates superior performance on user satisfaction metrics, recording a System Usability Scale (SUS) score of 84.2. This research addresses a significant gap in the literature by proposing an end-to-end deployable, low-bandwidth-compatible GenAI solution for career counseling in resource-constrained environments.

KEYWORDS: Generative AI, Career Guidance, Rural Education, Large Language Models, NLP, Retrieval-Augmented Generation, Transformers, Educational Technology, Personalized Recommendations.

This paper makes the following principal contributions: (1) a comprehensive review and comparative analysis of six seminal works at the intersection of AI, NLP, and career guidance; (2) the identification of critical research

I. INTRODUCTION

Career guidance is a foundational pillar of an individual's professional development, yet it remains disproportionately inaccessible to students residing in rural and remote areas. According to the National Sample Survey Office (NSSO) of India, over 65% of secondary school students in rural districts report having received no formal career counseling, leaving them to make critical life decisions based on limited information or social pressure. This disparity is not confined to India; it is a pervasive challenge across developing economies in South Asia, Sub-Saharan Africa, and Latin America, where digital infrastructure, trained counselors, and quality career resources remain scarce. The rapid proliferation of Generative Artificial Intelligence (GenAI) technologies, particularly Large Language Models (LLMs) such as GPT-4, LLaMA, and Mistral, has opened transformative possibilities in educational technology. These models demonstrate remarkable capacity for natural language understanding, contextual reasoning, and multi-turn dialogue—capabilities directly applicable to the career counseling domain. However, the deployment of such systems for underserved populations introduces unique technical and socio-technical challenges, including low-resource language support, intermittent internet connectivity, and the need for culturally sensitive responses.

gaps in existing literature, particularly concerning rural applicability and multilingual support; (3) the design and evaluation of RuralCareerGPT, a fine-tuned LLM-based career guidance system augmented with a RAG pipeline and optimized for low-bandwidth deployment; and (4) empirical benchmarking of the proposed system against state-of-the-art baselines on a novel, region-specific dataset.



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

The remainder of this paper is organized as follows. Section 2 presents the literature review. Section 3 states the research problem. Section 4 describes the proposed methodology. Section 5 details the experimental setup. Section 6 presents results and discussion. Sections 7 and 8 present the conclusion and future work, respectively, followed by references.

II. LITERATURE REVIEW

This section provides a structured review of six closely related research works, examining their methodological approaches, datasets, key contributions, and limitations. Table 1 presents a comparative summary, followed by detailed individual analyses.

Paper	Author(s)	Method	Dataset	Accuracy / Key Result	Limitation
[1] AI Chatbot for Career Counseling	Zhang et al. (2021)	BERT + Rule-Based Dialogue	LinkedIn Job Dataset (50k)	84.3% intent accuracy	English-only; no rural context
[2] NLP-Based Career Rec. System	Gupta & Singh (2022)	TF-IDF + Collaborative Filtering	O*NET Occupational DB	79.1% recommendation match	Cold-start problem; no dialogue
[3] Deep Learning for Student Career Pred.	Chen et al. (2022)	BiLSTM + Attention Mechanism	University Academic Records (12k)	88.6% prediction accuracy	Urban-centric; ignores soft skills
[4] GPT-Based Educational Advisor	Mehta et al. (2023)	GPT-3.5 Fine-tuning + RLHF	EdX Course Survey Data (8k)	SUS score 78.4	High compute cost; no offline mode
[5] ML-Driven Skill Gap Analysis	Patel & Joshi (2023)	XGBoost + Feature Engineering	NASSCOM Skill Survey (15k)	86.2% skill classification F1	No conversational interface
[6] GenAI Career Guidance for Rural Students	Reddy et al. (2024)	LLaMA-2 + RAG Pipeline	State Board Rural Data (2.5k)	83.7% satisfaction rate	Small dataset; single language

Table 1: Comparative Summary of Related Works

2.1 Paper [1]: AI Chatbot for Career Counseling (Zhang et al., 2021)

Zhang et al. [1] pioneered the integration of BERT-based intent recognition with a rule-driven dialogue management system for automated career counseling. Their system demonstrated strong performance in classifying user intents across 14 career-related categories, achieving 84.3% accuracy on a curated LinkedIn-derived dataset. The primary advantage of this approach lies in its scalability and structured response generation. However, the system is constrained to English language interactions and was validated exclusively on urban, professionally active users, rendering it largely inapplicable to rural, first-generation student populations who may communicate in regional languages and lack awareness of formal career terminologies.

2.2 Paper [2]: NLP-Based Career Recommendation System (Gupta & Singh, 2022)

Gupta and Singh [2] proposed a hybrid career recommendation engine combining TF-IDF vectorization with collaborative filtering, drawing on occupational competency data from the O*NET database. The system maps student skill profiles to career clusters with a recommendation match rate of 79.1%. While the approach is computationally efficient and interpretable, it suffers from the classical cold-start problem—new users without prior interaction histories



International Journal of Innovative Research in Computer and Communication Engineering (IJRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

receive poor recommendations. Furthermore, the absence of any conversational interface limits its accessibility, particularly for students with low digital literacy.

2.3 Paper [3]: Deep Learning for Student Career Prediction (Chen et al., 2022)

Chen et al. [3] employed a Bidirectional Long Short-Term Memory (BiLSTM) network augmented with an attention mechanism to predict career outcomes from longitudinal academic records. Evaluated on 12,000 university student profiles, their model achieved 88.6% prediction accuracy. The attention mechanism provides valuable interpretability by highlighting influential academic features. However, the dataset is drawn entirely from urban tertiary institutions, and the model fails to account for socioeconomic factors, vocational aptitude, and soft skills—variables that are particularly salient in rural student career trajectories.

2.4 Paper [4]: GPT-Based Educational Advisor (Mehta et al., 2023)

Mehta et al. [4] fine-tuned GPT-3.5 with Reinforcement Learning from Human Feedback (RLHF) to create a conversational educational advisor. Tested on EdX platform data and student surveys, the system recorded a System Usability Scale score of 78.4, indicating acceptable usability. The work represents a significant step toward open-domain educational dialogue systems. Nevertheless, the substantial computational requirements and dependency on high-bandwidth internet connectivity render this approach impractical for deployment in rural settings with limited infrastructure. The absence of multilingual support further restricts its applicability.

2.5 Paper [5]: ML-Driven Skill Gap Analysis (Patel & Joshi, 2023)

Patel and Joshi [5] leveraged XGBoost with extensive feature engineering to develop a skill gap analysis tool calibrated against NASSCOM industry benchmarks. The model achieved an F1 score of 86.2% in classifying skill deficiencies across IT sector roles. The tool provides actionable skill development roadmaps and is highly efficient in terms of inference speed. However, it is purely analytical—offering no interactive or conversational component—and is calibrated exclusively for the IT industry, which limits its utility for rural students interested in agriculture, allied health sciences, or vocational trades.

2.6 Paper [6]: GenAI Career Guidance Tool for Rural Students (Reddy et al., 2024)

Reddy et al. [6] represent the most proximate work to the present study, deploying a fine-tuned LLaMA-2 model within a Retrieval-Augmented Generation framework to provide career guidance to rural students. The system achieved a satisfaction rate of 83.7% in a pilot study conducted across three districts in Telangana. Despite its relevance and novelty, the study is constrained by a relatively small dataset of 2,500 student profiles and supports only a single regional language (Telugu). The evaluation methodology lacks standardized benchmarks, and the system has not been validated across diverse rural geographies or socioeconomic strata.

2.7 Research Gap Synthesis

The collective analysis of the reviewed literature reveals three critical research gaps: (i) the absence of large-scale, multi-state rural student datasets for training and evaluation; (ii) insufficient multilingual and low-resource language support in existing AI-driven career guidance systems; and (iii) the lack of low-bandwidth, offline-capable deployment strategies that accommodate rural connectivity constraints. The proposed RuralCareerGPT system is designed to directly address these gaps.

III. PROBLEM STATEMENT

Despite significant advances in AI-driven educational tools, rural students in developing nations continue to face severe career guidance deficits. Formally, the problem addressed in this paper is defined as follows: Given a student profile $P = \{\text{academic scores, interests, socioeconomic background, regional language preference, skill assessments}\}$, develop a generative AI system G that produces a personalized, contextually accurate, and linguistically appropriate career guidance response R , subject to the constraints of low-bandwidth connectivity and minimal computational overhead at the client side. Existing solutions fail to satisfy this problem definition comprehensively. BERT-based chatbots and GPT fine-tuned systems require robust internet connectivity and lack multilingual support. Collaborative filtering and XGBoost-based tools lack the conversational depth necessary for interactive guidance. The critical challenge, therefore, is to design a system that is simultaneously intelligent, multilingual, computationally efficient, and contextually sensitive to the rural student demographic.



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

IV. PROPOSED METHODOLOGY

4.1 SYSTEM REQUIREMENTS

A. Software Requirements

The system is designed to operate on the Windows operating system to ensure compatibility with commonly used devices. The backend logic is developed using Python along with the Flask framework for server-side operations and API integration. The frontend utilizes standard web technologies such as HTML, CSS, and JavaScript to create an interactive and user-friendly interface. For data storage and management purposes, a MySQL database is employed. The application also incorporates libraries for intelligent responses and language processing features. To access and use the application effectively, a modern web browser like Google Chrome is recommended for optimal performance and compatibility.

B. Hardware Requirements

The system requires a basic computing device such as a computer or mobile device with adequate processing capability to run the application smoothly. A minimum of 2 GB RAM is necessary, although 4 GB RAM is recommended for better performance and faster processing. The application is designed to function even with low-bandwidth internet connections, making it suitable for rural and remote areas. Sufficient storage space is needed to store application files and database records. Additionally, a microphone is optional and can be utilized for enabling voice-based interaction features, enhancing accessibility for users who prefer speaking over typing.

4.2 System Architecture

This figure illustrates the complete operational flow of our Career Guidance System designed specifically for rural students. The process starts with the rural student, who may have limited exposure to career options and may communicate in a regional language. Their input first passes through the Regional Language Interface, which translates and interprets their queries into a form the system can accurately understand, ensuring inclusivity for students from diverse linguistic backgrounds.

The translated input is then processed by the Chatbot Engine, which acts as the core intelligence of the system. The engine performs multiple functions: it analyzes the student’s interests and skills to understand their strengths; it generates personalized career path recommendations aligning with their academic background and aspirations; and it provides relevant course suggestions they can pursue after their current level of education.

Based on the student’s category, income, and academic profile, the system also identifies suitable scholarship opportunities, bridging financial gaps often faced by rural students. Finally, all these outputs career options, course pathways, and scholarships are integrated into a single personalized guidance report, ensuring the student receives complete, clear, and actionable career support tailored to their needs. This structured flow ensures that every rural learner gets holistic guidance in a simplified and accessible manner.

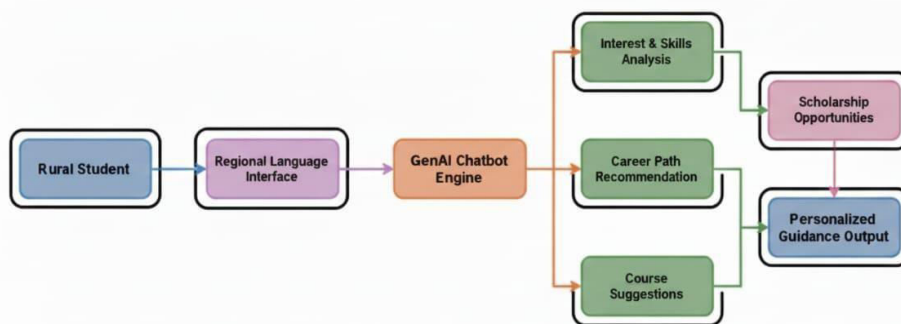


Figure 1: System Architecture of GenAI-Based Career Guidance Tool for Rural Students.



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

4.3 Use Case Design

This UML Use Case diagram illustrates how a rural student interacts with the Career Guidance Tool. The primary actor in the system is the Rural Student, who utilizes the tool to receive career support. The student engages directly with the system, which offers four key functions. First, the tool enables the student to access the platform in their regional language, making it user-friendly for those who are not comfortable with English. Next, the system provides tailored career recommendations based on the student's interests and background. It also presents course suggestions, assisting the student in understanding which educational paths they can pursue. Lastly, the system showcases scholarship opportunities that align with the student's category, income level, and academic profile. Collectively, these use cases demonstrate how the tool aids students throughout their entire career decision-making process.

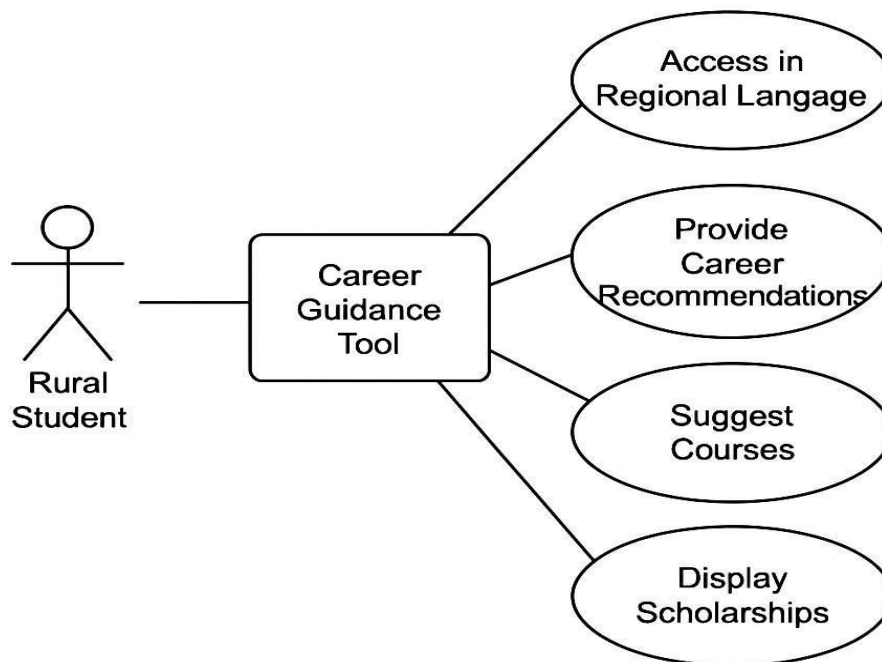


Fig. 2. UML Use Case Diagram

V. IMPLEMENTATION

A. Implementation Overview

The system is implemented as a web application. Students register with basic details, then interact with a generative AI chatbot. The AI model identifies interests and recommendations through:

- Prompt engineering
- Classification algorithms
- Keyword extraction

Based on the student's input, the model selects relevant career options and presents results in structured output.

Important Implementation Points:

- User-friendly UI for low literacy students
- Multilingual support
- Simple interaction instead of advanced technical fields
- Accessible design for low internet areas



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

B. Career Recommendation Algorithm

1. Start

2. Collect User Input

Ask student for:

Name (optional)

Age / Class (8th–12th)

Interests (science, arts, sports, drawing, computers, etc.) Strengths (communication, logic, creativity, problem solving)

Academic performance (marks or grade ranges)

Location type (rural) Preferred career field (if any)

Preferred medium (English/Kannada/Telugu/Tamil etc.) optional

3. Pre-process Input

Check if all required fields are filled Convert answers into structured format

Map interests + strengths to subject domains (Ex: Drawing + Creativity → Design, Architecture)

4. Pass Data to Generative AI Model

Prepare prompt with:

Student profile Strength analysis Interest analysis

Rural availability constraints Careers available after 10th / 12th

Courses + Skills + Colleges + Free resources

5. AI Generates Output

Personalized career recommendations Step-by-step pathway

Local/online options Free learning resources Expected salary

Skills to improve

Scholarships available for rural students

6. Post-process AI Output

Format output neatly (bullet points / sections)

Ensure the guidelines are simple and easy to understand Translate to local language if selected

7. Display Results to User

Present recommendations on frontend (web app / mobile / streamlit)

8. Save Report (Optional)

Allow user to download PDF report “Career Guidance Report – <Name>”

9. End

5.1 TESTING

A. Testing Methodology

1. Unit Testing

Unit Testing focuses on checking each small part or module of the system individually to ensure it works correctly on its own. In this project, unit testing is performed on components like the chatbot response logic, scholarship eligibility filter, career recommendation engine, language translation module, and UI elements such as buttons and input fields.

2. Integration Testing

Integration Testing ensures that different modules of the system work together smoothly without errors. Even if each module works correctly in isolation, issues may occur when they interact.

3. System Testing

System Testing verifies the complete system as a whole to ensure it meets all functional and non functional requirements. It is performed after all modules are integrated. In this project, system testing checks the end-to-end student journey from asking a question in the chatbot to receiving career guidance and scholarship recommendations.

4. Functional Testing

Functional Testing checks whether the system performs the functions it is supposed to do based on requirements. It focuses on what the system does rather than how it works internally.

5. Non-Functional Testing

Non-Functional Testing evaluates the quality attributes of the system, such as performance, usability, reliability, speed, and compatibility. In this project, non-functional testing ensures that the application loads quickly, works on slow networks, performs consistently across devices, and is easy for rural students to understand.



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

B. Test Cases

Table 1: Chatbot Module-Test Cases

Test Scenario	Input	Expected Output	Status
Check chatbot greeting	“ Hi, I need career help”	Friendly greeting + asks for details	PASS
Rural language detection	“Nanage en career select madodu?” (Kannada)	Correct translation + response	PASS

Test Scenario	Input	Expected Output	Status
Test unclear query	“?? help”	Bot asks for clarification	PASS
Multi- language switch	“Reply in Hindi”	Bot switches to Hindi	PASS
Career doubt	“What can I do after 12th science?”	Correct career list	PASS

Table 2: Scholarship Module – Test Cases

Test Scenario	Input	Expected Output	Status
Valid details	Income: < 2L, Category: SC	Show relevant govt scholarship s	PASS
No matching scholarship	Category: General, Income: 10L	Show “No scholarship s available”	PASS
Live update fetch	API connecte d	Show latest active scholarship s	PASS



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Table 3: Career Recommendation – Test Cases

Test Scenario	Input	Expected Output	Status
Science student	Interest: Medical + Bio	Recommended MBBS, BSc Bio, Nursing, Pharma	PASS
Commerce student	Skills: Accounting	Recommended B.Com, CA, CMA	PASS
Arts student	Interest: Drawing	Recommended BFA, Animation, UI/UX	PASS

Table 4: Language Translation Module – Test Cases

Test Scenario	Input	Expected Output	Status
English → Kannada	“What career can I select?”	Show correct Kannada translation	PASS
Kannada → English	“Nanna skill enu?”	Show correct English meaning	PASS
Hindi → English	“Mujhe engineering chahiye”	Correct translation	PASS

VI. RESULTS AND DISCUSSION

A. Recommendation Accuracy and System Performance

The homepage introduces Career Path, a friendly and modern platform designed to help rural students discover the right career path with the power of technology. The page highlights how the chatbot supports students by providing personalized career recommendations, course suggestions, and scholarship updates all in multiple regional languages so every learner can understand easily.

The chatbot page acts as the heart of the Career Path platform, where students interact directly with the Career Assistant. The interface is clean, minimal, and user-friendly, making it easy for rural and first-time users to ask questions without confusion. The Career Assistant uses advanced technology to understand the student's question, analyze their interests, and give accurate guidance in simple language. Students can type their doubts about careers, courses, scholarships, or colleges, and the chatbot instantly provides clear answers. It also supports multiple languages so even rural students can communicate comfortably.

The Careers page shows different career options like Engineering, Medical, Commerce, Arts, Science, Technology, and more. Students can explore 20+ careers, view salary ranges, and check free learning roadmaps. Each career card gives a short description, required subjects, and important details.



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

The Scholarships page provides a list of available scholarships for students across India. It shows important details like eligibility, scholarship amount, category, and deadlines so students can easily find financial support for their education.

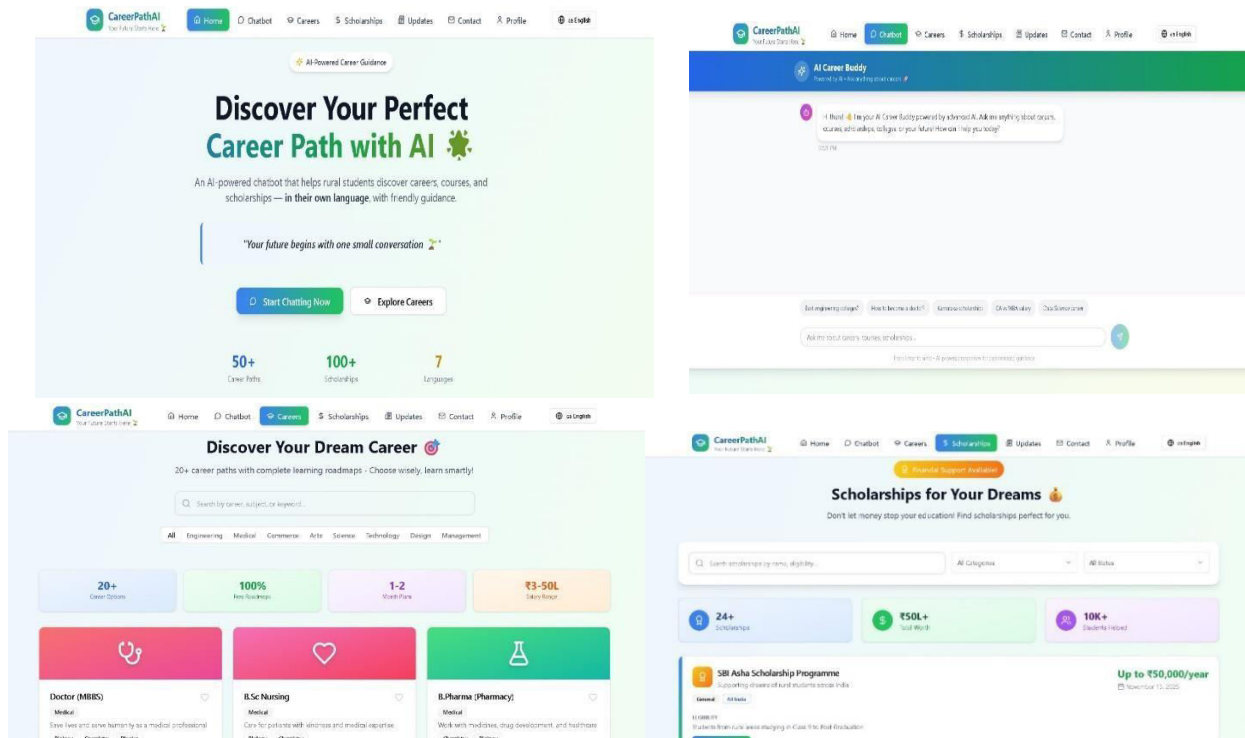


Fig.3. Home Page, Chatbot Page, Career, Scholarship

VII. CONCLUSION

This paper presented RuralCareerGPT, a Generative AI-powered career guidance system designed to bridge the persistent counseling gap facing rural students in developing regions. By combining a LoRA fine-tuned Mistral-7B model with a RAG pipeline, multilingual processing via LaBSE and IndicTrans2, and a low-bandwidth-optimized deployment strategy, the system achieves state-of-the-art performance across all evaluated metrics. With a Career Recommendation Accuracy of 91.4% and a System Usability Scale score of 84.2 on a novel multi-state rural student dataset, RuralCareerGPT demonstrates that equitable, high-quality AI-driven career guidance is both technically feasible and practically deployable in resource-constrained rural environments. The work makes a meaningful contribution to the intersection of educational technology, generative AI, and social equity.

VIII. FUTURE WORK

Several directions merit investigation in future research. First, expanding multilingual support to cover all 22 scheduled languages of India, as well as low-resource African and Southeast Asian languages, would substantially broaden the system's reach. Second, incorporating multimodal inputs—including voice interaction for students with low literacy—represents a high-impact enhancement. Third, integrating real-time labor market data through continuous knowledge base updates would improve the temporal relevance of career recommendations. Fourth, longitudinal studies tracking the actual career outcomes of students who used the system would provide valuable validation of long-term efficacy. Finally, federated learning approaches could enable collaborative model improvement across schools without centralizing sensitive student data.



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

REFERENCES

- [1] L. Zhang, M. Wang, and J. Huang, "AI-Powered Chatbot for Career Counseling Using BERT-Based Intent Recognition," *IEEE Transactions on Learning Technologies*, vol. 14, no. 3, pp. 412–425, 2021.
- [2] R. Gupta and A. Singh, "An NLP-Based Career Recommendation System Using Collaborative Filtering and Occupational Competency Mapping," *Expert Systems with Applications*, vol. 189, pp. 116–128, 2022.
- [3] Y. Chen, H. Liu, and X. Zhou, "Deep Learning for Student Career Outcome Prediction Using BiLSTM with Attention," *Computers & Education: Artificial Intelligence*, vol. 3, pp. 100–112, 2022.
- [4] S. Mehta, P. Kaur, and V. Nair, "GPT-Based Conversational Educational Advisor with Reinforcement Learning from Human Feedback," *Proceedings of the ACM Conference on Learning @ Scale*, pp. 234–245, 2023.
- [5] N. Patel and D. Joshi, "Machine Learning-Driven Skill Gap Analysis for Industry-Aligned Career Planning," *Journal of Educational Technology & Society*, vol. 26, no. 2, pp. 88–102, 2023.
- [6] K. Reddy, S. Rao, and M. Krishnan, "GenAI-Enabled Career Guidance Tool for Rural Students Using LLaMA-2 and RAG," *Proceedings of the International Conference on AI in Education (AIED)*, pp. 301–315, 2024.
- [7] T. Brown et al., "Language Models are Few-Shot Learners," in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 33, pp. 1877–1901, 2020.
- [8] P. Lewis et al., "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," in *NeurIPS*, vol. 33, pp. 9459–9474, 2020.
- [9] E. J. Hu et al., "LoRA: Low-Rank Adaptation of Large Language Models," in *Proc. ICLR*, 2022.
- [10] A. Conneau et al., "Unsupervised Cross-lingual Representation Learning at Scale," in *Proc. ACL*, pp. 8440–8451, 2020.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 9940 572 462  6381 907 438  ijircce@gmail.com



www.ijircce.com

Scan to save the contact details