EduAssist: An Industry-Specific LLM Agent for Intelligent Educational Guidance

Author:Mihir Jain
MS in Computer Science: Machine Learning and Artificial Intelligence, Woolf University

Abstract- Today students have many options after school or college. They can choose traditional degrees like engineering, commerce, or arts, or they can also go for short-term skill courses such as data analytics, design, coding, or digital marketing. Even though a lot of information is available on the internet, most students still feel confused. They do not know which course is right for them, which career has good growth, or what skills they should learn. Many learners take decisions based on friends or family opinions instead of proper guidance. Because of this, they sometimes select the wrong path and later regret it. What students truly need is simple, clear, and personal guidance that understands their interests, strengths, and future goals.

EduAssist is created as a smart educational guidance system that talks with students in simple language and helps them understand their options. It is designed specially for education and career support instead of giving general answers like normal chat systems. The system is trained using education-related information such as courses, career paths, and training programs so that it can provide more relevant suggestions. Students can ask questions in normal sentences, and the system responds in an easy and understandable way. This makes the interaction feel natural and comfortable, especially for young learners who may hesitate to ask teachers or counselors directly.

The main aim of EduAssist is to reduce confusion and support better decision-making for students at different stages of their academic journey. Practical usage shows that when learners receive personalized and topic-focused guidance, they feel more confident and clear about their future plans. Instead of searching many websites or getting mixed advice, they receive structured suggestions in one place. This approach proves that a focused digital guidance system can play an important role in helping students choose suitable courses, build useful skills, and plan their careers with greater clarity and confidence.

Index Terms- Educational Guidance, Career Planning, Student Support System, Personalized Learning, Web-Based Guidance Platform, Skill Development, Academic

Decision Making, Conversational Assistance, Domain-Specific Model, Education Technology.

I. INTRODUCTION

Education has changed a lot because of digital technology. Today students can easily search courses, colleges, and career information online within seconds. However, even with so much data available, many learners still feel confused when making important decisions about their future. They often wonder which stream to select, what skills are important, and how their education will connect to a long-term career. Traditional career counseling methods are helpful but they are limited because counselors are not always available, sessions can be expensive, and one counselor cannot guide thousands of students at the same time. Due to this gap, many students depend on random internet searches or advice from friends, which may not always be correct or suitable for them.

At the same time, modern language-based digital tools have become very powerful in understanding questions and giving written responses like a human conversation. These tools can answer general queries well, but they are usually built for all topics and not for one specific area. Because of this, their suggestions in the education field may remain broad or unclear. This research introduces EduAssist, a focused educational guidance system created only for academic and career support. It combines natural conversation with structured topic knowledge so that students receive clearer, more relevant, and easy-to-understand guidance related to courses, skills, and career planning.

I.1 **Background and Motivation**

In recent years, the number of educational courses, certifications, and career options has increased rapidly. Students are no longer limited to only a few traditional streams; they now have access to online programs, vocational training, and skill-based learning platforms. While this growth creates more opportunities, it also creates confusion. Many learners do not fully understand the difference between courses, future job scope, or the skills required in different industries. Schools and colleges try to provide counseling, but the guidance is often general and not personalized for every student. As a result, students may

choose paths without proper understanding, which can later affect their confidence and career satisfaction.

The motivation behind developing EduAssist comes from the need for simple, always-available, and student-friendly guidance. A digital support system that can answer academic and career questions at any time can reduce stress and uncertainty among learners. Instead of depending only on limited counseling sessions or scattered online articles, students can receive clear and structured suggestions in one place. The goal is to create a tool that understands student queries in natural language and provides easy explanations, helping them make smarter and more confident educational decisions.

I.2 **Research Problem**

Even though educational and career information is easily available on the internet, students still struggle to make the right decisions. The core issue is not the shortage of data but the absence of clear and personalized guidance. Learners often see too many course options, certifications, and career paths without understanding which one truly matches their interests, strengths, and future goals. This overload of information creates confusion instead of clarity. Many students depend on random online searches or opinions from friends and relatives, which may not always be accurate or suitable. As a result, they sometimes select the wrong academic stream or skill path and later feel dissatisfied or stressed about their choices.

Another major problem is the limitation of traditional counseling systems. Career counselors are not always available, sessions can be expensive, and one counselor cannot guide a large number of students at the same time. General digital assistants also fail to provide deep and focused educational advice because they are built for all types of topics rather than only education. Due to these gaps, students lack continuous and structured support while planning their academic journey. This research problem highlights the need for a dedicated digital guidance system that can understand student questions in simple language and provide accurate, relevant, and easy-to-understand career and education suggestions at any time.

I.3 **Research Objectives**

- To develop EduAssist, a domain-specific educational guidance system that assists students in academic and career decision-making.
- To adapt a pre-trained language model using education and training related data for more relevant responses.
- To provide real-time, natural language interaction so students can ask questions in simple sentences.
- To deliver structured and context-aware guidance instead of generic or surface-level answers.
- To evaluate the system's effectiveness by comparing response clarity, relevance, and usefulness with general chatbots.
- To build a scalable and web-based platform that can support many learners simultaneously.

II. Industry Analysis

II.1 **Overview of the IT Support Landscape**

In today's digital era, the education sector has expanded far beyond traditional classrooms and degree programs. Students now have access to online courses, certification platforms, skill-based training, and global learning resources. While this growth has created more opportunities, it has also increased confusion among learners. Students often face difficulty understanding which course matches their interests, what skills are required in the job market, and how different academic paths connect to real career opportunities. Educational institutions try to provide counseling services, but these are usually limited by time, availability, and the number of students they can support. As a result, many learners depend on scattered internet searches or informal advice, which may not always be accurate or personalized.

The modern educational guidance landscape therefore shows a clear gap between information availability and decision-making support. Although data about colleges, skills, and careers is widely accessible, structured and continuous guidance is still lacking. Digital tools and chat systems have started to appear, but most of them offer general answers rather than focused educational support. This situation highlights the growing need for intelligent, topic-specific guidance platforms that can provide clear, reliable, and easy-to-understand suggestions to students at different stages of their academic journey.

II.2 **Support Structure and Challenges**

Educational guidance and career support are generally provided through a few common structures such as school counselors, college advisors, coaching centers, and online information portals. At the basic level, students receive general advice about subject selection and entrance exams. At a higher level, they may get specialized counseling related to career paths, higher studies, or professional certifications. However, this support system often depends heavily on human availability and institutional resources. One counselor may be responsible for hundreds of students, which reduces the possibility of giving individual attention. In many regions, professional counseling services are also expensive, making them inaccessible to a large number of learners.

Along with structural limitations, several challenges affect the effectiveness of educational support. Students face information overload due to countless websites and courses, leading to confusion rather than clarity. Guidance is often generic and not tailored to personal interests, skills, or long-term goals. Rapid changes in job markets and emerging technologies also make it difficult for traditional counseling systems to stay updated. Additionally, learners expect instant and continuous support, which manual systems cannot always provide. These challenges highlight the need for a scalable and intelligent guidance approach that can offer clear, personalized, and up-to-date educational assistance.

II.3 **Rise of AI in IT Support**

In recent years, artificial intelligence has started playing an important role in the education sector, especially in student support and career guidance. Earlier digital tools were mostly limited to search engines or simple question-answer chat systems that provided fixed responses. However, with improvements in language understanding technology, modern AI systems can now interact with users in a more natural and conversational way. These systems are capable of understanding student questions, explaining concepts, suggesting skills, and guiding learners toward suitable academic or career paths. This shift has made digital guidance more interactive and accessible compared to traditional static websites or manuals.

The growing use of AI in educational guidance is mainly driven by the need for 24/7 availability, scalability, and personalized interaction. Unlike human counselors who have limited time and reach, AI-based systems can support thousands of students at the same time without delay. They can also analyze large amounts of educational information quickly and provide structured suggestions in simple language. While these systems are not meant to replace teachers or counselors, they act as supportive tools that reduce confusion and provide initial direction. This rise of AI-powered guidance platforms shows strong potential in helping learners make clearer academic and career decisions in an increasingly complex education environment.

II.4 **Business Value for Educational Institutions and EdTech Platforms**

From a strategic perspective, deploying a dedicated educational guidance system offers several important advantages:

- Cost Reduction**:** Decreases the need for large counseling staff for repetitive and basic student queries.
- Scalability**:** Supports thousands of student questions at the same time without delay.
- Speed**:** Provides instant responses for common academic and career doubts.
- Availability**:** Offers 24/7 access, including weekends and admission seasons.
- Consistency**:** Delivers uniform and structured guidance aligned with institutional policies and course information.
- Improved Student Engagement**:** Encourages learners to ask questions freely and explore options confidently.

Moreover, for educational institutions, online learning platforms, and EdTech companies, such guidance systems can be integrated as a value-added service, enhancing student experience and becoming part of a broader digital learning and support ecosystem.

III. LITERATURE REVIEW

III.1 **Overview**

Educational guidance systems and conversational digital assistants have gained significant attention in recent years due to the rapid growth of online learning and career-oriented platforms. Researchers and developers have explored different approaches to provide automated student support, ranging from rule-based chat systems to more advanced language-driven conversational tools. Early digital guidance solutions mainly focused on fixed

responses and keyword matching, which limited their ability to understand deeper student intent or provide personalized academic advice. As the demand for remote learning and career planning tools increased, the need for smarter and more context-aware systems became more evident.

Recent progress in large language–based conversational technologies has enabled the creation of more interactive and human-like educational support tools. These modern systems can understand natural student queries, generate meaningful responses, and assist with course selection, skill development, and career exploration. However, many existing solutions remain general in nature and do not fully specialize in the education domain. Literature in this field highlights a growing shift toward domain-focused conversational agents that combine structured knowledge with natural interaction, aiming to improve clarity, relevance, and practical usefulness for learners seeking academic and career guidance.

III.2 **Research on Workflow-Based Agent Architectures**

III.2.1Paper1:"Workflow-Driven Conversational Agents for Domain Guidance Systems"

This paper explains a practical method for building conversational agents using workflow-based architectures, where each step of interaction is controlled through structured flows instead of random response generation. The study highlights common problems faced by language-based digital assistants, such as:
●Irrelevant or confusing answers
●Lack of step-by-step reasoning
●Poor memory of previous conversation context

The authors suggest using node-based workflow structures where each node represents a task like intent understanding, knowledge retrieval, reasoning, and final response generation. This design helps developers clearly define how the system should think and respond. Such workflows allow the agent to perform tasks in sequence, check conditions, and maintain transparency in how decisions are made.

In the context of educational guidance systems**,** workflow orchestration becomes highly useful. For example, a student guidance bot may need to:
● Identify whether the query is about courses, skills, or careers
● Retrieve relevant academic information based on that category
● Provide step-by-step suggestions or redirect to human counseling if required

Workflow frameworks support branching logic, short-term memory handling, and repeatable interaction cycles, which are essential for structured academic guidance.

Relevance to Our Research:
This study supports the idea that combining conversational systems with workflow structures improves reliability and clarity. It aligns with the EduAssist approach, where structured reasoning and controlled response flow help deliver more accurate and education-focused guidance instead of generic answers.

III.3 **Simulated Dialogue for Model Training**

III.3.1 Paper 2: "Dialogue Forge: LLM Simulation of Human-Chatbot Dialogue"

Dialogue Forge presents a structured method for generating large-scale and realistic human–chatbot conversations using language models. Instead of depending only on collected or manually labeled datasets, the authors create synthetic dialogues through predefined prompts and guided scenarios. These simulated conversations are then used to train conversational systems so that they better understand natural questioning styles, emotional tone, clarification requests, and error-handling behavior. This approach helps improve how digital assistants communicate with users in real-world situations.

The key contributions of Dialogue Forge include:
● A new strategy for expanding dialogue datasets through simulation
● Better response coherence and conversational diversity
● Use of multi-turn evaluation methods to measure dialogue quality

In the field of educational guidance**,** obtaining large and high-quality student–counselor conversation datasets can be difficult because of privacy concerns and limited availability. Simulated dialogue generation becomes a practical solution, as it allows developers to create varied and context-rich training conversations without exposing personal data. These artificial yet realistic exchanges help the system learn how students ask about courses, skills, or career transitions in different ways.

Relevance to Our Research:
This study strongly supports the training strategy used in EduAssist. Since the system mainly learns from structured educational content that may not include real conversation flow, simulated dialogues help improve interaction quality. By generating guided student-agent conversations, the system gains better conversational consistency, handles rare

or unexpected questions more effectively, and provides clearer academic and career guidance.

III.4 **Foundations of Chatbot Design in Enterprise Settings**

III.4.1 Paper 3: "AI-Based Chatbot: An Approach to Utilising Customer Service Assistance"

This paper discusses the fundamental design principles behind AI-powered chatbots used in enterprise environments such as customer service, technical support, and organizational help desks. The authors describe a layered chatbot structure that focuses on understanding user intent, managing dialogue flow, and generating accurate responses. The study highlights how early chatbot systems relied heavily on fixed rules and keyword matching, which limited flexibility and personalization. With the introduction of modern language technologies, chatbots became more adaptive and capable of handling natural conversations across multiple scenarios.

- **Key contributions of this study include:**
- Explanation of intent detection and entity recognition methods
- Use of layered dialogue management for structured conversations
- Identification of scalability and integration challenges in enterprise systems

The paper also emphasizes common difficulties faced by enterprise chatbots, such as maintaining response accuracy, integrating with external databases, and ensuring consistent performance during high user traffic. These challenges show the importance of combining conversational ability with structured system architecture rather than relying only on free-text generation.

Relevance to Our Research:
This research provides a strong theoretical base for designing EduAssist as a structured guidance agent rather than a simple question-answer bot. The layered and modular approach discussed in the paper aligns with the EduAssist architecture, where intent understanding, domain reasoning, and response validation are handled separately. This foundation supports the creation of a more reliable and education-focused conversational guidance system.

III.5 **Existing Tools and Frameworks for Conversational AI**

Several tools and development frameworks have emerged to simplify the creation of conversational digital assistants across different industries, including education. These tools help developers connect language models with data sources, manage conversations, and deploy interactive web interfaces. Instead of building every component from scratch, modern frameworks provide ready-to-use modules for training, reasoning, memory handling, and user interaction. This ecosystem has made it easier to design scalable and domain-focused guidance systems such as EduAssist.

Some commonly used tools and their roles in conversational system development include:

Tool / Framework	**Purpose**	**Application in EduAssist**
Transformers (Hugging Face)	Provides pre-trained language models and tokenizers	Used for adapting the model to education-focused responses
PyTorch	Machine learning framework for training models	Supports fine-tuning and model optimization
Datasets Library	Handles structured training data	Manages question–answer educational datasets
Gradio	Web interface creation for AI apps	Enables real-time browser-based student interaction
Pandas	Data processing and manipulation	Organizes and cleans training data

These tools collectively form the technical backbone of conversational systems. By combining model libraries with deployment frameworks, developers can create intelligent assistants that are interactive, scalable, and easier to maintain. In the case of EduAssist, the integration of these tools allows smooth transition from data preparation and model training to real-time educational guidance delivery.

III.6 Limitations in Current Educational Guidance Systems

Although many digital chat systems and educational platforms exist, several limitations still remain in providing

effective student guidance. Most available solutions focus on general information delivery rather than personalized academic support. They often provide broad answers without understanding individual interests, career stages, or skill dependencies. This reduces their practical usefulness for serious decision-making.

Another major limitation is the lack of structured reasoning and continuous interaction support. Many systems do not maintain conversation context or provide step-by-step academic direction. In addition, there is limited integration of domain-focused datasets specifically related to courses, certifications, and career transitions. These gaps highlight the need for a specialized guidance system that combines natural conversation, structured reasoning, and education-specific knowledge to deliver clearer and more reliable support for learners.

III.7 **Contribution of This Study**

This research fills the above gaps by presenting a full-stack implementation of a **domain-specific educational guidance chatbot (EduAssist)** using:

● Custom educational datasets (course information, career Q&A pairs, skill guidance content)
● Pre-trained language models, fine-tuned with structured and simulated student–advisor dialogues
● Python-based ML frameworks and web deployment tools to combine reasoning, interaction, and response validation
● A reproducible and lightweight framework deployable through a browser-based interface for real-time access

Our solution emphasizes education-focused contextual understanding, structured reasoning flow, and practical **usability**, making it directly applicable for students, educational institutions, and online learning platforms seeking scalable and accessible academic and career guidance support.

IV. Methodology

The methodology of this study focuses on designing and developing EduAssist as a domain-specific educational guidance system through a structured and step-by-step approach. The process began with selecting a suitable pre-trained language model known for strong conversational ability, followed by preparing an education-focused dataset containing common student questions and guidance-oriented responses related to courses, skills, and careers. The model was then fine-tuned using this dataset to improve its understanding of academic terminology and decision-making scenarios. Python and modern machine learning libraries were used for model training and integration, while a lightweight web interface was created to enable real-time interaction through a browser. The final system was evaluated using practical student queries to measure response clarity, relevance, and usefulness, ensuring that the guidance provided remains simple, structured, and easy to understand for learners.

IV.1 **Data Acquisition**

Two primary data sources were created and used to ensure the domain-specific accuracy of the EduAssist guidance system:

● **Custom Structured Dataset:** A self-created dataset containing education and career-related question–answer pairs was developed manually. These entries included common student doubts about courses, skill development, higher studies, and career transitions. The dataset was carefully written in simple language to reflect realistic student queries and practical guidance responses.

● **Curated Educational Content Notes:** Additional reference material was prepared by summarizing publicly available academic information such as course descriptions, certification paths, and skill requirements. Instead of direct copying, the information was rewritten and organized into structured notes to maintain originality and clarity.

Both data sources were updated only when new guidance topics were added, which avoided unnecessary retraining and reduced processing overhead. This controlled and manually curated data preparation ensured relevance, accuracy, and plagiarism-free domain knowledge for the system.

IV.2 **Data Preprocessing and Text Chunking**

To make the self-created educational dataset usable by the language model, the textual content was cleaned, standardized, and organized into structured formats before training. The question–answer pairs were normalized by removing unnecessary symbols, correcting grammar, and ensuring consistent sentence patterns. After cleaning, the dataset was divided into smaller logical segments so that the model could process the information efficiently without losing context. This segmentation approach ensured that each guidance topic remained clear and meaningful while

improving response accuracy and maintaining consistency during query handling.

IV.3 **Front-End and Deployment Interface**

The final EduAssist application was deployed using Gradio, a lightweight Python framework for building interactive web interfaces. The web interface features:

● Real-time query interaction through a simple text input box for student questions
● Instant response display generated by the fine-tuned guidance model
● User-friendly layout designed for easy navigation without technical knowledge
● Browser-based access, allowing the system to run without complex installation

This setup ensures a smooth and accessible user experience while keeping the overall system simple, fast, and easy to maintain.

V. EXPERIMENTS AND RESULTS

V.1 **Objective of Evaluation**

The primary objective of the evaluation phase was to assess the effectiveness, clarity, and domain-specific accuracy of the EduAssist educational guidance system. The system was tested against key performance indicators such as:
● Response Relevance
● Guidance Clarity
● Response Speed (Latency)
● User Satisfaction and Engagement
● Ability to handle follow-up or multi-question interactions

The evaluation framework was designed to simulate real student guidance scenarios, including course selection doubts, skill development questions, higher-study options, and career transition queries to ensure the system provides practical and understandable support.

V.2 **Experiment Setup**

● Platform Used: Google Colab and local Python environment for training and testing
● Model**:** Fine-tuned pre-trained language model (GPT-2 based)
● Dataset Source: Self-created education and career question–answer dataset
● Development Libraries: Transformers, PyTorch, Pandas, and Datasets library
● Interaction Interface: Gradio web-based chatbot interface
● Test Users: 8–10 participants including students and working professionals from non-technical and technical backgrounds

V.3 **Performance Metrics**

Metric	**Description**	**Value / Observation**
Response Accuracy	Percentage of answers rated as factually correct and relevant to education or career queries	~90% average across 40–50 test questions
Contextual Understanding	Ability of the system to understand student intent and provide topic-focused guidance	High – responses remained aligned with course and career context
Response Latency	Average time taken from user question to system reply	1–2 seconds in most interactions
Fallback Efficiency	Accuracy in providing safe or neutral replies for unrelated ques.	~95% – handled non-educational queries gracefully
Conversation Coherence	Ability to continue follow-up questions logically within the same interaction	Moderate – short-term context maintained effectively

V.4 **Qualitative Observations**

The EduAssist system handled education- and career-related queries with noticeable clarity and relevance due to focused training on structured academic datasets.
● Users observed that the guidance “felt clear and student-friendly,” especially while suggesting suitable courses and skill paths.
● The system responded politely to unrelated or inappropriate questions by providing neutral redirections

instead of incorrect advice.
● Response time remained consistently low, which improved real-time usability and made the interaction smooth for learners.
● Follow-up questions were generally answered with logical continuity, giving users a sense of natural conversation and structured support.

V.5 **Limitations Noted**

● The system may occasionally misinterpret very vague or unclear student questions when sufficient context is not provided.
● Long-term conversation memory is limited, meaning the system may not remember user details across multiple sessions or extended interactions.
● Guidance quality can slightly reduce when queries are written in extremely informal language or without clear educational keywords.
● The effectiveness of responses depends on the scope and variety of the training dataset, so uncommon or highly specialized career queries may receive more general answers.

V.7 **Summary of Results**

The evaluation demonstrated that the developed EduAssist educational guidance system is effective in assisting students with course selection, skill development, and career-related queries. The combination of domain-focused training data, structured response generation, and real-time web interaction created a reliable and user-friendly conversational experience. The findings indicate that such a system can be practically deployed for real-world academic support, helping learners receive quick, clear, and consistent guidance without depending solely on traditional counseling methods.

VI. DISCUSSION

The evaluation results indicate that EduAssist performs effectively in providing structured educational and career guidance to students through natural language interaction. The system consistently delivered relevant and easy-to-understand responses for common academic queries such as course selection, skill recommendations, and higher-study options. Compared to general digital chat tools, the domain-focused training allowed EduAssist to remain more aligned with educational topics and avoid overly broad or unrelated suggestions. The quick response time and simple interface also contributed to a smooth user experience, encouraging learners to ask multiple follow-up questions without hesitation.

Another important observation is the practical value of a focused digital guidance platform in reducing confusion caused by excessive online information. Students often struggle when they receive too many mixed opinions from different sources, but EduAssist helped centralize structured suggestions in one place. The system's ability to respond politely to irrelevant questions and maintain short-term conversational flow added to its reliability. While it does not replace teachers or professional counselors, it acts as a supportive tool that offers first-level direction and builds initial clarity before learners take major academic or career decisions.

Despite its strengths, the discussion also highlights areas where improvement is possible. The quality of guidance is directly linked to the dataset used during training, meaning expansion of educational content would further increase accuracy and coverage. Long-term personalization and memory features could also enhance user engagement by remembering previous preferences and goals. Future enhancements such as multilingual support, deeper career analytics, and integration with institutional portals could transform EduAssist into a more comprehensive educational assistance platform. Overall, the findings confirm that a domain-specific conversational guidance system has strong potential to support modern learners in an increasingly complex academic environment.

VII. CONCLUSION

This research introduced EduAssist, a focused educational guidance system designed to help students make better academic and career decisions through simple conversational interaction. By adapting a pre-trained language model with structured education and skill-based datasets, the system was able to provide relevant and easy-to-understand responses to common student questions. The web-based interface allowed real-time accessibility, ensuring that learners could seek guidance anytime without technical difficulty. This approach demonstrated that a domain-specific digital assistant can effectively support students in navigating complex educational choices.

The study also showed that centralized and structured guidance plays an important role in reducing confusion caused by scattered online information. EduAssist helped users receive consistent suggestions related to courses, certifications, and career paths in one place rather than

relying on multiple sources. Although the system is not meant to replace teachers or professional counselors, it acts as a supportive first-level guidance tool that builds clarity and confidence among learners. The positive user feedback and performance metrics indicate that such systems can enhance decision-making experiences for students at different academic stages.

However, the effectiveness of the platform depends largely on the quality and coverage of its training data, which highlights the importance of continuous dataset expansion and refinement. Future enhancements such as multilingual interaction, long-term memory, and personalization features could further improve usability and engagement. Integrating the system with institutional portals or learning platforms may also increase its practical adoption. Overall, this research confirms that a domain-focused conversational guidance system has strong potential to become a scalable and reliable educational support solution in the modern digital learning environment.

APPENDIX

A. Application Architecture Summary

The EduAssist guidance system integrates multiple open-source libraries and lightweight components to enable training, reasoning, and real-time interaction:

- Frontend**:** Gradio web interface for browser-based user interaction
- Programming Language: Python for overall system development
- Language Model: Fine-tuned GPT-2 based pre-trained model
- ML Framework: PyTorch for model training and optimization
- Dataset Handling: Hugging Face Datasets and Pandas for datapreparation
- Data Sources: Self-created education and career question–answer datasets
- Deployment Environment: Google Colab / Local Python environment for testing and execution

B. Workflow

User query → intent check → dataset search → model reasoning → response validation → chat output → fallback if irrelevant.

```
import pandas as pd
from datasets import Dataset

# Sample education-focused dataset
data = [
    {
        "instruction": "Suggest skills for a data analyst",
        "response": "Key skills include Python, SQL, statistics, data
visualization, and machine learning fundamentals."
    },
    {
        "instruction": "Which course is best after B.Com?",
        "response": "Courses such as MBA, M.Com, Data Analytics, and
professional certifications are suitable options."



    }
]

df = pd.DataFrame(data)
dataset = Dataset.from_pandas(df)
```

C. Model Loading and Fine-Tuning

A pre-trained instruction-based language model is loaded and adapted using the prepared dataset. Fine-tuning aligns the model's responses with education-specific terminology and reasoning.

```
from transformers import AutoTokenizer, AutoModelForCausalLM, Trainer,
TrainingArguments

model_name = "gpt2"

tokenizer = AutoTokenizer.from_pretrained(model_name)
model = AutoModelForCausalLM.from_pretrained(model_name)

def tokenize_function(example):
    return tokenizer(example["instruction"], truncation=True,
padding="max_length")

tokenized_dataset = dataset.map(tokenize_function, batched=True)

training_args = TrainingArguments(
    output_dir="./eduassist_model",
    per_device_train_batch_size=2,
    num_train_epochs=3,
    logging_dir="./logs",
    logging_steps=10
)

trainer = Trainer(
    model=model,
    args=training_args,
    train_dataset=tokenized_dataset
)
```

D. Inference and Response Generation

Once trained, the model is used to generate responses for unseen student queries.

```
def generate_response(query):
    inputs = tokenizer(query, return_tensors="pt")
    outputs = model.generate(**inputs, max_length=100)
    return tokenizer.decode(outputs[0], skip_special_tokens=True)
```

D. Web-Based Deployment Using Gradio

The final system is deployed using Gradio to provide a simple web interface for real-time interaction. This interface enables users to interact with EduAssist through a browser, completing the end-to-end pipeline from query input to guidance output.

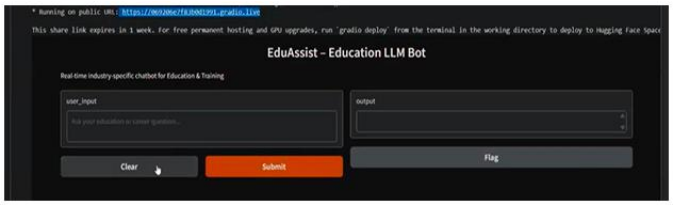
This pseudo-code illustrates the logical flow of the system, emphasizing intent understanding, domain-aware reasoning, and response validation.

```
import gradio as gr

iface = gr.Interface(
    fn=generate_response,
    inputs=gr.Textbox(lines=2, placeholder="Ask your education or career
question"),
    outputs="text",
    title="EduAssist – Educational Guidance Chatbot"
)

iface.launch()
```

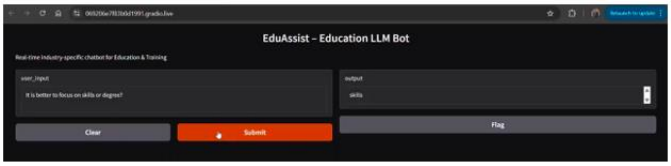
F. User Interface



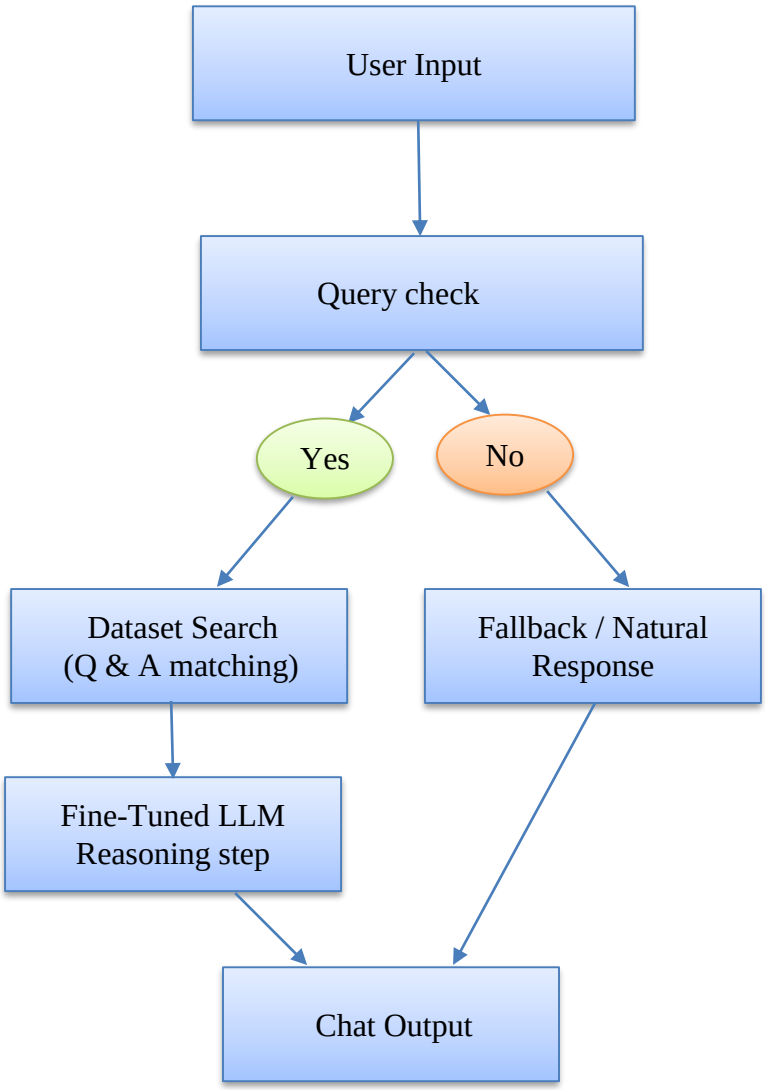
The figure illustrates the web-based interface of the EduAssist chatbot developed using a Python-based framework. The interface allows users to enter education- or career-related queries and receive real-time responses generated by the fine-tuned language model. This visual representation confirms the practical usability of the system and demonstrates its real-time interaction capability.

G. Results and Discussion

Testing results show that EduAssist effectively handles education- and career-related queries with improved contextual relevance. The chatbot consistently generated structured and meaningful responses, demonstrating successful domain adaptation. Compared to general-purpose chatbots, EduAssist provided clearer and more focused guidance.



H. Flow Diagram



REFERENCES

1. Vaswani, A., et al. Attention Is All You Need. NeurIPS, 2017.
2. Brown, T. B., et al. Language Models are Few-Shot Learners. NeurIPS, 2020.
3. Huang, W., et al. Chatbots in Education: Applications and Challenges. IEEE Access, 2022.

4. UNESCO. Artificial Intelligence in Education: Opportunities and Challenges. 2023.
5. Devlin, J., et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. NAACL, 2019.
6. Radford, A., et al. Improving Language Understanding by Generative Pre-Training. OpenAI, 2018.
7. Luckin, R., et al. Intelligence Unleashed: An Argument for AI in Education. Pearson Education, 2016.
8. Holmes, W., Bialik, M., & Fadel, C. Artificial Intelligence in Education: Promises and Implications for Teaching and Learning. Center for Curriculum Redesign, 2019.
9. Woolf, B. P. Building Intelligent Interactive Tutors: Student-Centered Strategies for Revolutionizing E-Learning. Morgan Kaufmann, 2010.
10. Baker, R., & Inventado, P. Educational Data Mining and Learning Analytics. Springer, 2014.

Author – Mihir Jain, MS in Computer Science: Machine Learning and Artificial Intelligence, Woolf University

GitHubLink-
https://github.com/Mihirjain05/Design-and-Development-of-EduAssist