iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** XI **Month of publication:** November 2024

DOI: https://doi.org/10.22214/ijraset.2024.65069

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue XI Nov 2024- Available at www.ijraset.com

EduAssist: Chatbot for Student Admission Queries

Sahil Baviskar[1], Kshitija Deshmukh[2], Lokesh Deshmukh[3], Nishica Kothawade[4], Manisha Mali[5]
[1, 2, 3, 4, 5]*Department of Computer Engineering, Vishwakarma Institute of Information Technology, Pune*

Abstract: ***Across engineering institutes, inquiries from students and parents surge during the admission period, covering topics like admissions processes, eligibility, college details, fees, curriculums, scholarships, and placements. Current methods like calls, emails, and visits are inefficient, leading to delays and higher workloads for staff.***
An AI-powered chatbot using Retrieval-Augmented Generation (RAG) and NLP offers a solution by providing 24/7 instant, accurate, and personalized responses in multiple languages, including English, Hindi, and regional languages. It automates answers to common questions, enhancing access to key information, easing staff workload, and improving the overall user experience.
Keywords: Multilingual Support, FAQ Data Augmentation, Voice Recognition, RAG, NLP

I. INTRODUCTION

It is worth noting that during the application phase, the engineering institutes receive a lot of inquiries from potential students and their guardians. Most of these inquiries are often limited to the admission process, eligibility criteria, institution particulars, tuition fees, program curricula, financial aid ways, job prospects etc. The conventional mechanisms of answering them (such as telephone inquiries, emails, and site visits) do not work quite well, and delays in giving information often result in providing different information on the same topic and an increased workload for administrative personnel, which makes the experience user unfriendly and wastes the resources of the organization.

In response to these challenges, AI-based chatbots have been developed which have proven to be extremely effective. This is because of technologies such as Retrieval-Augmented Generation (RAG) and Natural Language Processing (NLP), the chatbots respond instantly, accurately and tailored to the user at any given time. Furthermore, they can conduct n number of conversations at the same time providing the same information with no time lags and errors as it is with humans.

One of the most significant benefits of this chatbot is its multilingual support, which includes Hindi, English and other regional languages to cater to the linguistic needs of diverse users. This feature makes it possible for a large section of the population to have access to the information regardless of the language that is used in the communication. Moreover, such chatbots can also provide automated responses to frequently asked questions, such as admission requirements, course fees, courses available and details on placements, which will considerably reduce the burden on administrative personnel.

As a result of integrating RAG and NLP methodologies, the chatbot is also able to understand the context of user queries and provide more relevant and personalized information. This understanding is crucial for ensuring that prospective students and their families enjoy a seamless and engaging experience especially in understanding the admissions process.

In conclusion, the integration of artificial intelligence-based robotic chat agents in engineering schools is a measured approach to increase communication and improve processes. For example, these chatbots reduce the workload for institutional personnel while improving user experience by offering 24/7 multilingual assistance and automated responses to common queries. The incorporation of AI-powered devices, for example, chatbots, into schools will be necessary for aiding workflows and enhancing productivity in today's technological world.

II. LITERATURE REVIEW

Gen AI helps you to create new content, videos, images. This project helps to create chatbot using LLM, NLP and Gen AI. Answer the query content provided through pdf. The responses are generated through RAG. The primary technologies utilized include Python libraries like LangChain, PyTorch for model training, and Hugging Face's Transformers library for accessing pre-trained models like Llama2, GPT- 3.5. Aims to develop chatbot with sensible response in the form of pdf.[1]

This paper represents Chatbot system in educational domain. It created to assist university student's queries. The goal to develop a model for communication and providing proper response to students. System develops to recognize queries and provide answers to students using AI and NLP. [2]

©IJRASET: All Rights are Reserved

Advancement in LLM have catalyzed the widely integration of chatbot across various field like education, finance, e-commerce, marketing and health-care. This chatbot are reshaping the landscape of consumer service and user interaction. This application uses technology like LLM and data gathered from college website to give responses like human discourse. The chatbot aims to reduce the time users spend navigating the complex college website by efficiently retrieving college-related information. [3]

A chatbot is computer program makes conversation with help of Artificial Intelligence (AI). Each time it takes input from user and generated response with its initial knowledge. Implement an online chatbot system to assist website users uses NLP technology. Eliminate manpower and available 24x7.[4]

This paper introduces a college enquiry chatbot that helps answer student questions quickly and efficiently. It uses natural language processing and machine learning to understand and respond to common queries. Tests showed that the chatbot provides accurate answers in a short time, saving time for both students and staff. The chatbot can interact with users through text and voice, making it more engaging. Python's Rasa X framework, a powerful tool for building chatbots, is used to develop the system. This study shows how useful chatbots can be in education and suggests there is room for more research in this field.[5]

III. PROPOSED METHODOLOGY

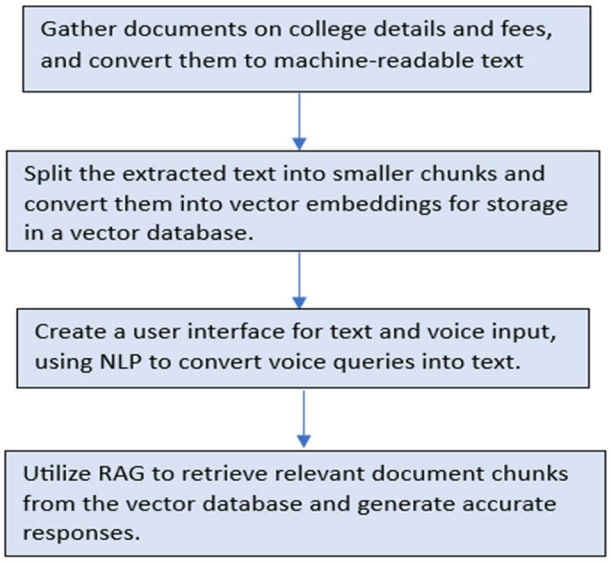


Fig.1: System Flow Diagram

The purpose of this project is to enhance information sharing in educational institutions by presenting a systematic procedure for developing an AI-based chatbot that caters to students' queries. The methodology comprises several critical steps:

A. Data Collection

The first step in this process is information gathering about the University. Such details include course availability, fee structures, job placement statistics, volunteering opportunities, awards, and other information. Most of this information comes from college official websites, brochures, and similar materials. It is also important to know that the information collected should be accurate, up to date and include all the details that the future student and their parent could ever wish for. It is done this way to help in the systematic data collection process that creates a wealth of information that helps build the knowledge base of the college chatbot.

B. Text Extraction

After the collection of the papers, the information has to be organized in such a manner that machines will be able to read it. Scanned papers use Optical Character Recognition (OCR) technology for text recognition and extraction from images. This is usually the case where documents are in images such as TIFF or furnished in a digital format such as MS word and pdf. This transformation is very important because it is the one that changes the form of the data, which the chatbot can process and use when interacting in real time.

C. Document Chunking

When the text is obtained, it is expected that the content will be subdivided into smaller and homogeneous portions. This approach called document chunking improves both the accuracy of data retrieval and the relevance of questions answered by the chatbot. The pieces could be made in logical divisions such as paragraphs or topics so that each one contains information that is interconnected to one another. For instance, all the details regarding scholarship may be put in one category. Since the chatbot has to answer to the questions posed by the users in a quick and accurate manner, the need for chunking becomes acute.

D. Vector Database Storage

After Processing, the text segments are stored in a vector database. Each chunk is converted into a vector embedding using tools like Sentence Transformers that encode the given text into meaningful numerical values. Thereafter, this method optimizes the efficiency of similarity searches. By encoding the information in a vector form, the system can quickly fetch and compare chunks based on their relevancy to the user information request, thereby greatly enhancing the speed of the system.

E. User Query Input

The design of the chatbot enables parents and students to pose their questions with ease considering the friendly UI. This interface can also reach a wider audience since it provides both a text input and voice input. This is one of the more important design goals; in other words, ease of use must allow for an easy way to navigate the system. This step is very important in defining the basic purpose of the communication, which is the interaction of users with the chatbot.

F. Query Processing

A system's first task upon receiving a request is to ascertain whether it is voice or textual input. If it is voice input, the spoken words are transcribed into written words with the help of Natural language processing techniques. This is necessary for the chatbot because it helps to read the input correctly and understand what the user intends to do in order to get the relevant information. This also enhances the NLP system integrated into the chatbot so that it can understand and respond to a wide range of queries.

G. Retrieving Relevant Documents

The chatbot employs the RAG (Retrieval-Augmented Generation) approach in completing the task by retrieving the most appropriate document chunks from the vector database after the user's query has been processed. This means finding chunks that are most similar to the query embedding so that the data returned is relevant to the user's question and also fitting for the purpose at hand. Because of the RAG system, informativeness is enhanced which allows for a more relevant response.

H. Generating Responses

Once relevant document chunks have been fetched, the system generates a reply by utilizing the information within these document chunks. This process involves synthesizing the fetched materials in order to form a coherent and informative answer that effectively resolves a query raised by the user. The information provided for end-users must be both relevant and clear; therefore, the response mechanism must concentrate on these aspects more than anything else. This last bit is very important because it helps determine the how satisfied the user is with the system in place considering the functionality of the chatbot.

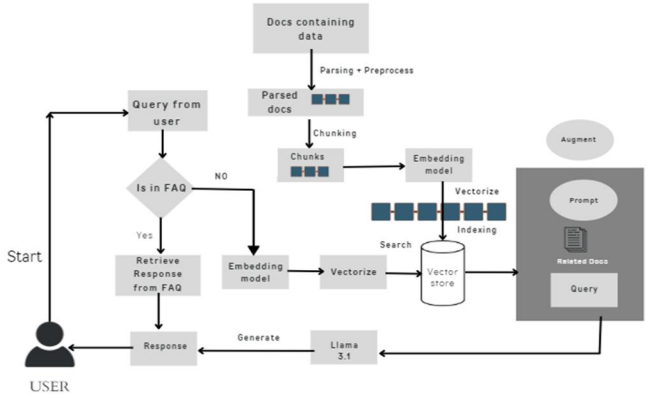


Fig.2: RAG ChatBot Architecture

IV. TECHNIQUE INVOLVED IN PROPOSED METHODOLOGY

A. Given Below are Some Techniques used for AI Chatbot:

1) LangChain Community: A framework for chatbots with document retrieval, text transformation and dialogue management components. It ensures the chatbot first extracts the necessary information from college-related documents (such as fees, placements etc.) before generating a response.

2) FAISS (Facebook Artificial Intelligence Similarity Search): is a Vector similarity search algorithm that helps in storing archival data subsets as a vector database. Therefore, the relevant document is retrieved almost immediately upon query submission by the student. In this process, the Long-Chain CharacterTextSplitter often divides the text data into smaller portions which is then converted into vectors using Hugging Face embedding and retrieval is performed by storing the vectors in FAISS.

3) LangChain Groq (ChatGroq): This API integration enables the chatbot to generate relevant and contextual responses after fetching the relevant information from the FAISS database.

4) NLTK (Natural Language Toolkit): NLTK is useful in cleaning up and carrying out text transformation, as in this case, documents become easier to analyze and interpret. It helps in parceling documents for an efficient FAISS storing process.

B. Techniques used for Data Processing are

1) Tesseract OCR: This technology utilizes is the ability of the ChatBot to pull any text from scanned images of college brochures, admission policies, and fee structures and convert the scanned pdf files into usable data.

2) UnstructuredPDFLoader (from LangChain): This makes it possible to interpret and pull-out important information from formulated PDFs. The information found in the PDFs such as the fees details, placements stats, etc. is uploaded in the system and divided in parts to assist the processing of the parameters for the chatbot.

3) Streamlit: It offers easy-to-use website where the learners can type in their questions. It guarantees user-friendliness by acting as the main interface for interacting with the chatbot.

C. *Machine Learning Models used are*

1) Hugging Face Transformers: These models are responsible for transforming query and document text into vector representation. The Chatbot uses the embeddings stored in FAISS to generate human-like responses.

2) Speech Recognition (Google Speech-to-Text): Converts voice-based student queries into text. If a student uses voice input, this tool helps translate the spoken question into a format that the chatbot can process and respond to.

V. RESULTS

The chatbot has achieved significant improvements in efficiency and user experience through the implementation of Retrieval Augmented Generation (RAG) and Natural Language Processing (NLP). By using RAG, the chatbot is highly efficient because it is trained specifically on the required data, ensuring accurate and relevant responses. Additionally, NLP has enabled multi-language support and speech-to-text capabilities, allowing students and parents to receive answers in their preferred language, enhancing accessibility. This has led to 90% of students communicating in their primary language and reduced workload of college staff by 50%.

Here is the following output after performing the proposed methodology:

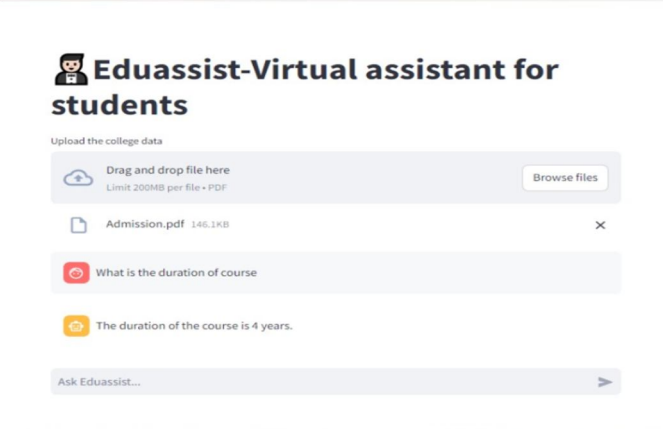


Fig 3.1: Document Upload

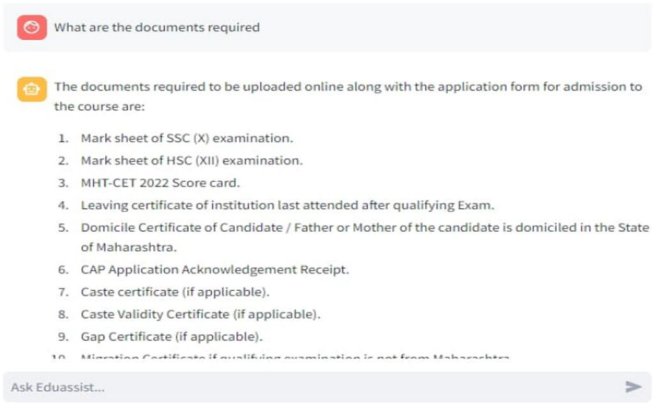


Fig 3.2: Student Query 1

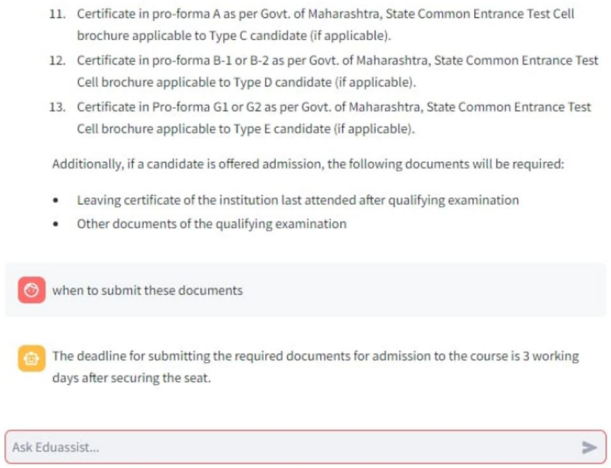


Fig3.3: Response with stored memory

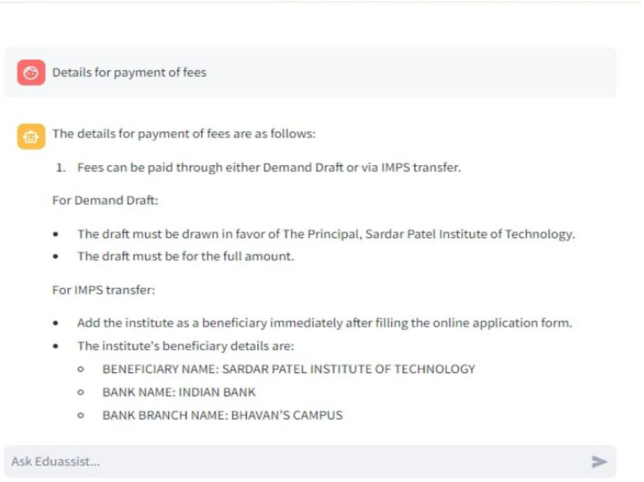


Fig 3.4: Student Query 2

VI. APPLICATIONS

1) Academic Queries: Anytime, students can get answers to their questions regarding the courses, syllabus, exam time table and grading system.

2) Administrative Assistance: The chatbot makes it easy for students to carry out the paperwork of the institution like admissions, fee payments, and submission of documents thus lightening the workload of the office personnel.

3) Career Counselling: The bot can counsel students on internships, placements, skill enhancement programs and assistance in governmental jobs for effective career guidance of the students.

4) Multi Language Support for Domestic and International Students: The use of a multilingual chatbot, can help students understand in their preferred language which is especially helpful for people who cannot speak English or Hindi.

VII. FUTURE SCOPE

The expansion of future chatbot development shall prioritize multilingual features which encourage more regional and international languages. Also, I believe that integrating further AI algorithms would probably help in developing individual student personalized recommendations. Furthermore, it may also be possible to improve the chatbot's functionality by including more advanced machine learning methods that will help it adapt through experience. Better speech recognition technologies may be looked into to also allow more natural and fluid interactions. Lastly, widening the scope of services that this digitally enabled assistant would provide to include exam schedules, interaction with alumni, and event related notifications among others would make this tool more beneficial to students and staff alike through providing a more enhanced virtual assistant.

VIII. CONCLUSION

The student enquiry chatbot emerges as a beneficial asset for schools and colleges in providing scalable and economical solutions to manage the students' queries. With the help of RAG and NLP, students are provided answers instantly in context. Other access-friendly features include multilingual and voice support. Future extensions may include increasing the language scope and enhancing the recommendation systems for personalization of the student interaction experience.

REFERENCES

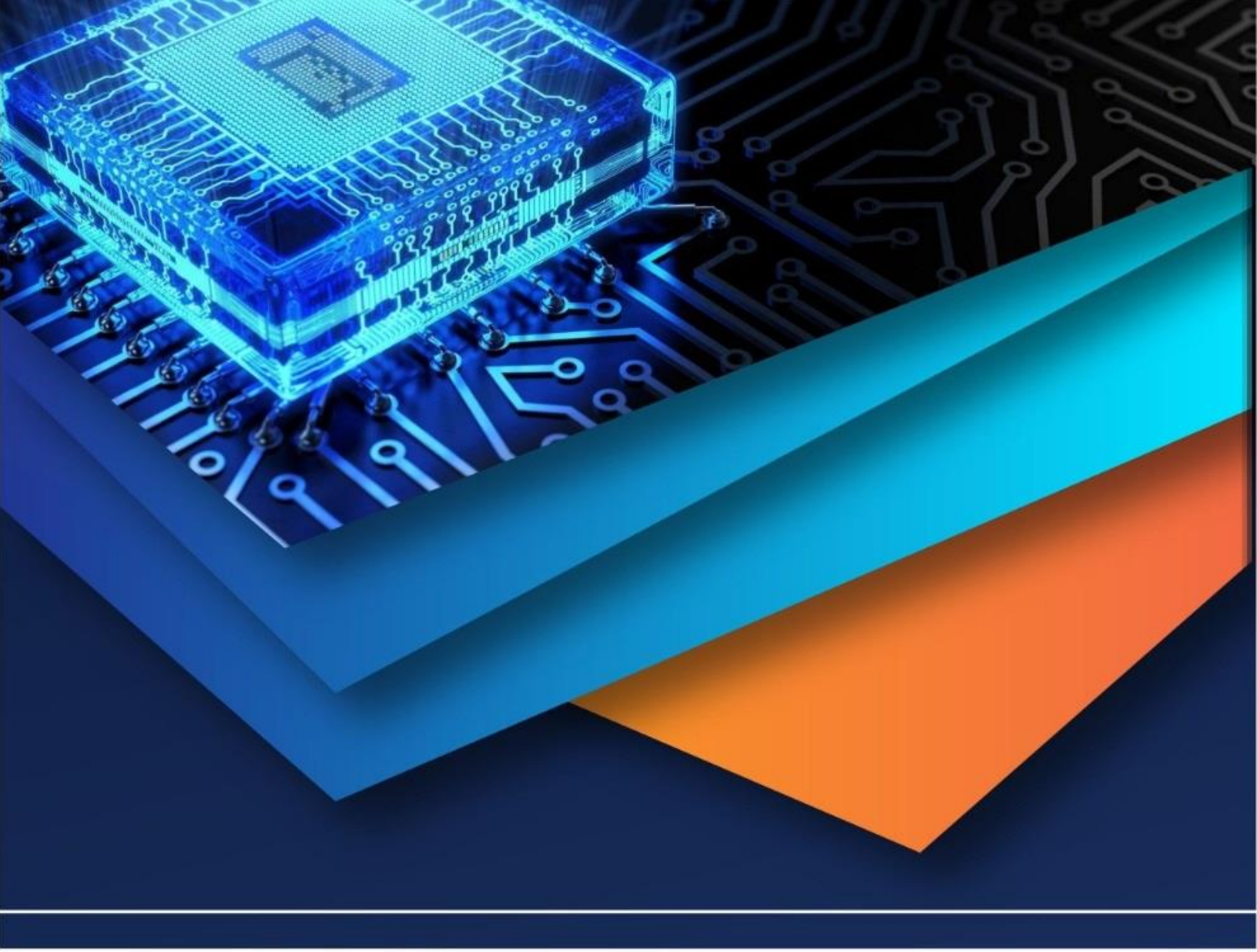
[1] Ananya G, Dr. Vanishree K. "RAG based Chatbot using LLMs", International Journal of Scientific Research in Engineering and Management (IJSREM), Volume: 08 Issue: 06, ISSN: 2582-3930, June–2024, DOI: 10.55041/IJSREM35600

[2] Walaa Hassan, Shereen elBohy, Mina Rafik, Ahmed Ashraf, Sherif Gorgui, Michael Emil, Karim Ali. "An Interactive Chatbot for College Enquiry", Journal of Computing and Communication, Vol.2, No.1, PP. 20-28, 2023.

[3] G.V. Pavan, K.M.N. Kumar, P. Bhargav Reddy, K.N.V.D.V.S. Aneesh, N. Gowthami, "SMART CAMPUS BOT: COLLEGE QUERY RESPONSE CHATBOT", International Research Journal of Modernization in Engineering Technology and Science, Volume:06/Issue:03/March-2024, e-ISSN: 2582-5208.

[4] V. Adarsh, B. Koushik, D. Mahesh. "CHATBOT USING NATURAL LANGUAGE PROCESS (NLP)", International Research Journal of M odernization in Engineering Technology and Science, Volume: 05/Issue:02/February-2023, e-ISSN: 2582-5208.

[5] Gokul R, Kalaivani S. "AI College Enquiry Chatbot System", Journal of Emerging Technologies and Innovative Research (JETIR), Volume 10, Issue 2, February 2023, (ISSN-2349-5162).



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)