

Contextual Conversational Intelligence : Leveraging Language Models For College Chatbot By Retrieval Augmented Generation

Anergha K Mohan¹, Thomas John², Nikhil P S³

^{1,2,3} *Computer Science and Engineering , Model Engineering College , Kochi , India*

Abstract - By harnessing the capabilities of Large Language Models (LLM), the proposed system evolves into a sophisticated conversational agent, capable of understanding nuanced user inputs, providing context-aware responses, and delivering an unparalleled level of engagement. This chatbot is also equipped to assess a student's likelihood of being accepted into a specific program at the college using a machine learning (ML) model. The model analyzes the student's rank and eligible quota (such as reserved categories) in conjunction with historical admission data to gauge their chances of acceptance. It also recommends alternative programs where the student may have a better chance of admission based on their rank. The chatbot also features a smart Applicant Tracking System (ATS) that assists students in evaluating their resumes. This system provides keyword matching suggestions based on job descriptions and generates an overview of the resume. Additionally, it highlights areas for improvement and highlights the resume's distinctiveness. The methodology involves the development and deployment of a website chat bot integrated with Lang Chain, with the potential of advanced language models in transforming website chat bot interactions using Retrieval augmented generation (RAG) technique to ground LLM to generate responses to user queries based on a custom knowledge-base.

Key Words: Chatbot, Generative AI, Large Language Model, Website, Machine Learning, Retrieval-augmented generation , Web scrapping , Resume , ATS.

1.INTRODUCTION

In the era of digital communication, chat bots have emerged as a powerful tools for enhancing user engagement and providing real-time support. The proposed system here, revolves around the integration of a cutting-edge chat bot powered by Large Language Models (LLM) into the website, aiming to revolutionize user interaction and enhance the overall website experience. By Leveraging LLM to endow the chat bot with advanced conversational abilities, allowing it to comprehend natural language queries and generate contextually relevant and coherent responses and hence providing Conversational Intelligence, Intent Recognition, Proactive Engagement. This chat bot system is designed to evaluate a student's prospects of gaining admission to a specific course within a college through a machine learning (ML) model. By considering factors such as the student's rank and eligible quotas (including reserved categories), the model utilizes

historical admissions data to calculate the likelihood of acceptance. Furthermore, it offers suggestions for alternative programs where the student may have improved chances of admission based on their ranking. The dynamic and personalized responses are anticipated to boost user engagement, encouraging longer sessions and increased exploration of website offerings, enhancing Natural Interaction, Enhanced User Engagement and User-Centric Adaptability.

It also features an intelligent Applicant Tracking System (ATS) that assists students in analyzing their resumes. This system provides keyword matching suggestions related to job descriptions and offers a comprehensive summary of the resume. Furthermore, it highlights areas for improvement and points out distinctive features that can make the resume more appealing to potential employers. This process utilizes the Retrieval-Augmented Generation technique to enhance the LLM, allowing it to generate responses to user queries based on a tailored knowledge base. Data is retrieved from the website through web scraping, ensuring that the chatbot's responses are grounded in relevant and up-to-date information.

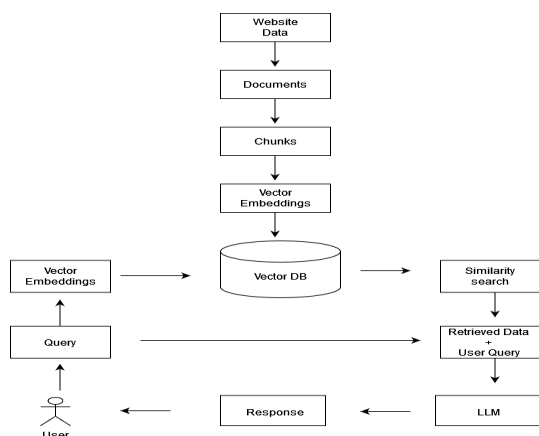
2.DATA AND SOURCES OF DATA

For the proposed system of RAG chat bot, custom website data is retrieved by web scrapping using Lang Chain Selenium URL Loader, enabling the collection of data from dynamic web pages. For admission likelihood prediction, historical institution specific data on past admission cut-off trends, quota-based admission statistics, course popularity over last 10 years is used.

3.PROPOSED SYSTEM AND DESIGN

The proposed system uses custom data from the institution's website by web scrapping, employing the LangChain Selenium URL Loader to facilitate the data retrieval process. Once the scraped data is gathered, the Lang Chain CharacterTextSplitter is employed to break down large pieces of text into manageable documents and chunks. By splitting the text into smaller pieces, it ensures the data can be efficiently processed. The Google Generative AI Embeddings component is used

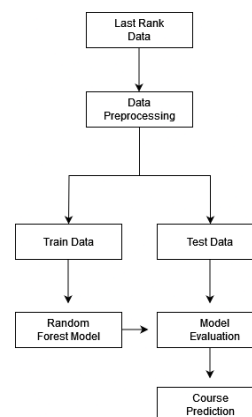
to convert chunks of text into embeddings -numerical representations of the text. These embeddings encode the semantic meaning of the text, enabling the model to better understand relationships between words, phrases, and documents. This is crucial for tasks like searching for relevant information, comparing different documents. The Google Gemini is used to generate responses based on user queries and the provided embeddings. It is responsible for the core interaction between the user and the chatbot, where users ask questions about institution , admission probability, or program recommendations. The LLM uses embeddings and other contextual information to produce coherent and relevant responses, enhancing the chatbot's ability to provide personalized insights. Course prediction is done, based on Reservation quota and last rank , using Ensemble learning technique combining Random Forest models. Resume analysis done by the prompt engineering techniques leveraging the functionalities of LLM.



The Random Forest Classifier, implemented through Scikit-learn, is a machine learning model used to predict the likelihood of a student gaining admission to a particular program. Random Forest is an ensemble learning method that builds multiple decision trees and merges them to increase the accuracy and stability of predictions. By training this model on historical admission data, student ranks, and quota eligibility, it predicts the probability of admission based on the provided input. This prediction is the key to the chatbots functionality, that helps to provide accurate admission insights.

Matplotlib and Seaborn are used for data visualization, helping present insights in a clear and understandable format about the historical data. These libraries are especially useful for visualizing the distribution of last-rank cutoff data, admission trends, and quota-specific admission patterns. For example, after making a prediction about a student's chances of admission, the chat bot could show a graph of historical cutoffs for the preferred course, helping the student better understand

their eligibility. These visualizations make it easier to grasp complex patterns in the data and improve the user experience.



4.RESULTS AND DISCUSSION

The workflow was successfully implemented using a combination of tools and techniques to achieve web scraping, data manipulation, embedding creation, model building, and visualization. The results from each step are summarized below.

MEC Bot

Enter your query:

tell about Research

Ask Me

```

{
  "query": "tell about Research"
  "result":
    "Govt. Model Engineering College is the first Engineering College in the state recognised by Cochin University of Science and Technology as the Research Centre of University in the branch of Electronics and Communication Engineering. The major areas of research are VLSI and Embedded Systems, Signal and Image processing, Opto Electronics, Biomedical Engineering etc.

    The Department of Mechanical Engineering is offering both full time and parttime PhD programmes in Energy Management, Renewable Energy, Reliability Engineering, and Thermal Sciences. The ongoing research works are in Reliability Engineering, Electric Spring, and Electro Chemistry and Fluid Flow.

    Recognized as a research centre of Cochin University of Science and Technology (CUSAT) in the year 2010, Model Engineering College has been promoting active research in the fields of Signal &Image Processing and Optical Communication. Later, under A P JAbdul Kalam Technological University (KTU), research has been
  
```

These depicts the responses based on the data from the institutions website.

MEC Bot

Enter your query:

tell about Academic Eligibility

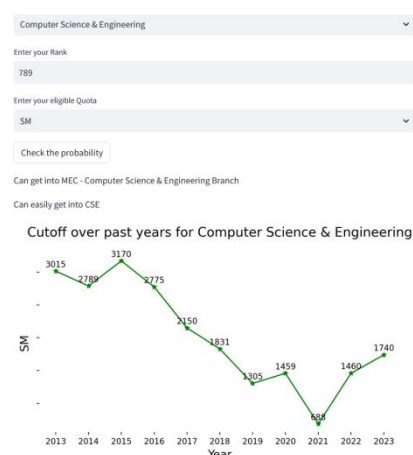
Ask Me

```

{
  "query": "tell about Academic Eligibility"
  "result":
    "Academic Eligibility: Candidates who have passed Higher Secondary Examination, Kerala, or Examinations recognized as equivalent thereto, with 50% marks in Mathematics separately and 50% marks in Mathematics, Physics and Chemistry / Bio-technology / Computer Science / Biology taken together are eligible for admission."
}
  
```

Below response shows the probability of getting admission into a particular program based on previous

year cut off ranks. Second response is based on the ML models prediction higher probable course.



This also shows the analysis of the students resume and its curated response.

Smart ATS

Improve Your Resume ATS

Paste the Job Description

Evaluate internal systems for efficiency, problems, and inaccuracies, and develop and maintain protocols for handling, processing, and cleaning data
Work directly with managers and users to gather requirements, provide status updates, and build relationships

Upload Your Resume

Drag and drop file here
Limit 200MB per file • PDF

test.pdf 207.8KB

Submit

JD Match: 75%

MissingKeywords: ["Machine Learning Algorithms", "Cloud Computing", "Big Data Analytics", "Agile Development", "Hadoop"]

MissingKeywords: ["Machine Learning Algorithms", "Cloud Computing", "Big Data Analytics", "Agile Development", "Hadoop"]

Summary: The resume demonstrates strong technical skills in software engineering, data science, and data analysis. The candidate has experience working with Python, SQL, and R, as well as cloud platforms such as AWS and Azure. The candidate also has a strong understanding of big data technologies such as Hadoop and Spark. However, the resume lacks some of the keywords that are commonly found in job descriptions for software engineers, data scientists, and data analysts. These keywords include machine learning algorithms, cloud computing, big data analytics, agile development, and Hadoop.

What makes this resume stand out from others: The resume stands out from others due to the candidate's strong technical skills and experience working with a variety of technologies. The candidate also has a strong understanding of big data technologies, which is in high demand in the current job market.

What improvements can be done: The candidate can improve their resume by adding some of the missing keywords to their resume. They can also add more detail to their work experience, such as describing the specific projects they worked on and the results they achieved. Additionally, the candidate can consider adding a section to their resume that highlights their skills in machine learning algorithms, cloud computing, big data analytics, agile development, and Hadoop.

5.CONCLUSION

The work focuses on enhancing the factuality of language model to give customised chat domain. It predicts the Course allotment for students based on previous year cut off ranks. The development of an AI and web-based interactive enquiry chatbot addresses challenges in obtaining timely and accurate information

about related events. The chatbot uses AI technologies and Natural Language Processing (NLP) to engage with users, offer suggestions, and provide quick access to even nuanced questions. This solution enhances user experience, saves time, and improves accessibility to related information.

REFERENCES

1. Lia Fauzia, Raden Budiarto Hadiprakoso, Girinoto, "Implementation of Chatbot on University Website Using RASA Framework", 2021 4th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)
2. Se-Hoon Lee, Ung-Hoe Lee, Ji-Woong Kim, Yeon-Su Noh, "College Admissions Counseling ChatBot based on a Large Language Models", 2023 Korea Computer and Information Society
3. Ori Ram, Yoav Levine, Itay Dalmedigos Dor Muhlgay Amnon Shashua Kevin Leyton-Brown Yoav Shoham, AI21 Labs, "In Context Retrieval-Augmented Language Models", arXiv:2302.00083v3 Aug 2023
4. Yasmina Al Ghadban, Huiqi (Yvonne) Lu, Uday Adavi, Ankita Sharma, Sridevi Gara, Neelanjana Das, Bhaskar Kumar, Renu John, Praveen Devarsetty, Jane E. Hirst, "Transforming Healthcare Education: Harnessing Large Language Models for Frontline Health Worker Capacity Building using Retrieval-Augmented Generation", medrxiv.org
5. Akari Asai, Zequi Wu, Yizhong Wang, Avirup Sil, Hannaneh Hajishirzi, "Self-rag: learning to retrieve, generate, and critique through self-reflection
6. Ayyoub Ajmi, University of Missouri - Kansas City, School of Law, "Revolutionizing Access to Justice: The Role of AI Powered Chatbots and Retrieval-Augmented Generation in Legal Self-Help", law.umkc.edu/faculty/works/949
7. Mandar Kulkarni, Praveen Tangarajan, Kyung Kim, Anusua Trivedi Flipkart Data Science Seattle, Washington, USA, "Reinforcement Learning for Optimizing RAG for Domain Chatbots", AAI 2024 Workshop on Synergy of Reinforcement Learning and Large Language Models
8. Ahmet Yusuf Alan, Enis Karaarslan*, Omer Aydin, "A RAG-based Question Answering System Proposal for Understanding Islam: MufassirQAS LLM", 2024.