# UniBuddy: Gen AI Assistant

Deepak Ayyasamy, Tarun Siga, Sai Karthik Naladala

UMKC School of Science and Engineering

## Introduction

Navigating the vast landscape of university options can be daunting for students seeking the ideal educational experience. UniBuddy emerges as a solution to this challenge by providing a user-friendly virtual assistant that consolidates all pertinent university information into a single, accessible platform. By leveraging advanced technologies and methodologies, UniBuddy aims to streamline the process of university exploration and decision-making, ultimately empowering students to make informed choices about their academic futures.

## Methodology

**Secure User Management:** Implementing robust authentication with Flask and Firebase ensures data privacy and integrity, offering users hassle-free registration.

**Real-time Communication:** Integration of Firebase with Streamlit enables instant access to information, facilitating seamless communication between users and the virtual assistant.

**Collaborative Application Development:** Utilizing Flask, Firebase Authentication, Firestore, and Streamlit for backend infrastructure and intuitive interface enhances user engagement within the UniBuddy platform.

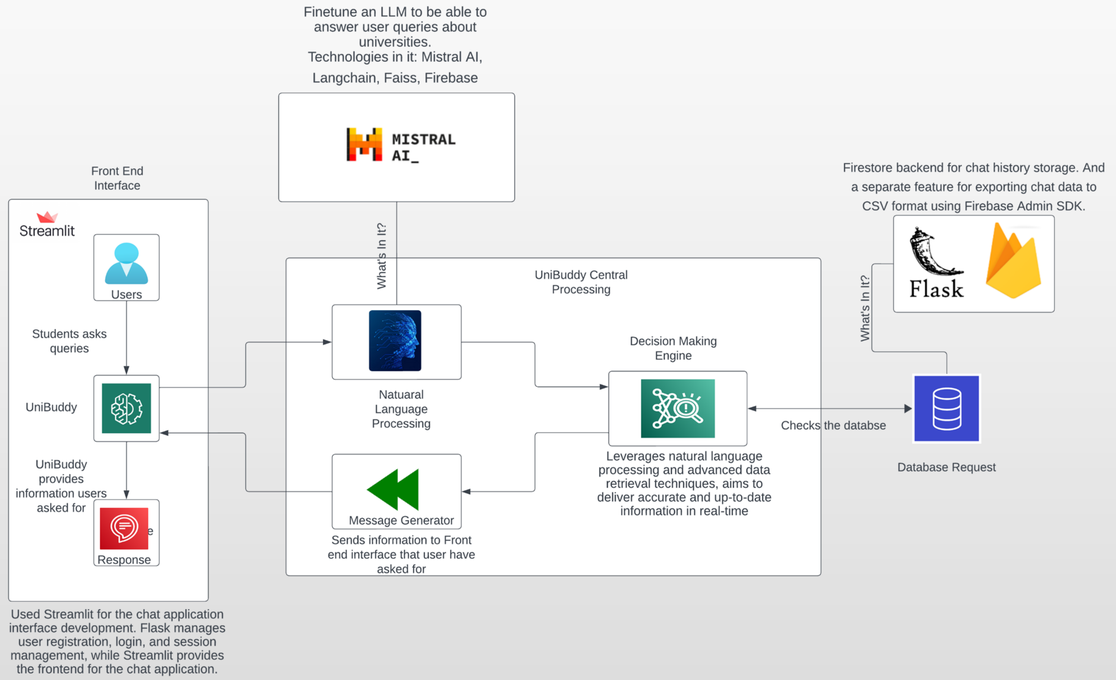
**Efficient Text Representation:** Text embeddings, created from text chunks using models like Word2Vec or BERT, enhance response accuracy and relevance, stored in a vector database for efficient retrieval.

**Advanced NLP Techniques:** Leveraging Mistral AI and LangChain, UniBuddy comprehends user queries and generates human-like responses, ensuring contextually relevant information in real-time.

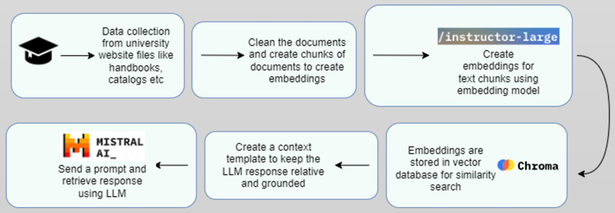
**Efficient Data Ingestion:** Developing an ingestion module with multithreading and multiprocessing techniques ensures seamless processing of large data volumes, providing timely information to users.

**Optimized GPT Module:** UniBuddy's local GPT module, optimized for question-answer tasks, integrates embeddings and vector stores for efficient text retrieval, ensuring a seamless user experience.

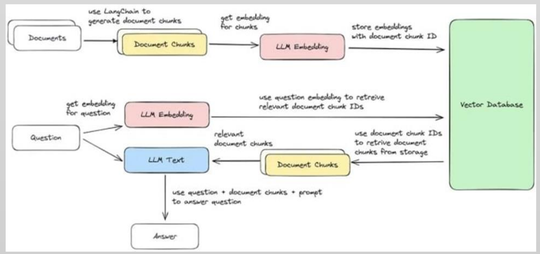
## Concept Diagram



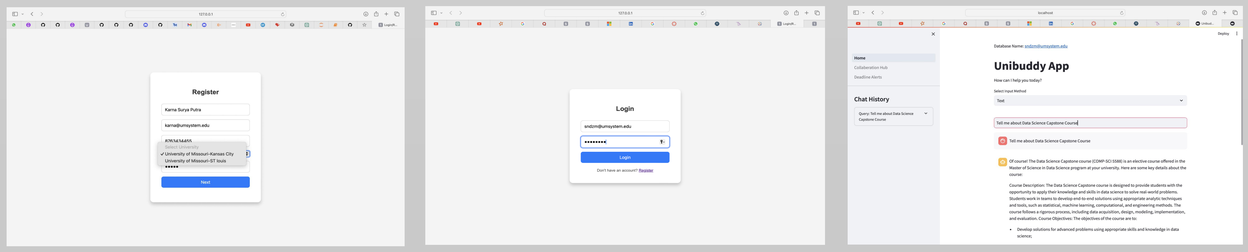

## Model Diagram




## Embedding Workflow




## Front End




## Evaluation

To evaluate the performance of our fine-tuned LLM module, we compared the results of our LLM model with the top Gen AI application in the market, like, GPT 3.5, Gemini 1.5 and Perplexity AI.

The evaluation included quantifying the responses of different LLM responses based on four factors -

- **Comprehensiveness:** Comprehensiveness refers to the extent to which a response covers all necessary aspects or components of the topic or query at hand.
- **Relevance:** Relevance pertains to how directly and closely a response addresses the specific topic, question, or query posed.
- **Clarity:** Clarity refers to how easily understandable and coherent a response is to the intended audience.
- **Helpfulness:** Helpfulness assesses the practical utility or value of a response in aiding the audience's understanding or addressing their needs.

We calculated this for 25 different questions and then arrived at mean values. Results suggested Unibuddy outperformed GPT 3.5 and Perplexity AI and came second to Gemini only by a bit.

| Aspect | UniBuddy | GPT 3.5 | Gemini | Perplexity |
|---|---|---|---|---|
| Reality | 10 | 10 | 10 | 6 |
| Comprehensivesness | 8 | 8 | 10 | 4 |
| Clarity | 10 | 8 | 10 | 6 |
| Helpfulness | 8 | 8 | 10 | 4 |
| Overall Score | 9 | 8.5 | 10 | 5 |

## Revenue Generation and Future work

- UniBuddy possesses significant potential for revenue generation through its evolution into a Software as a Service (SaaS) model. By offering subscription-based access, UniBuddy can provide users with comprehensive tools and resources to enhance their academic standing, career planning, and personality development.
- Converting UniBuddy into an user oriented SaaS product that can tailor responses based on the user preferences. Reduce response time of LLMs by using better processors.

## References

1. https://towardsdatascience.com/how-to-build-an-llm-from-scratch-8c477768f1f9
2. https://thomaschereckal.medium.com/how-to-create-your-own-llm-model-2598615a039a