

GenAI Project Report – HelpMate AI

Members:

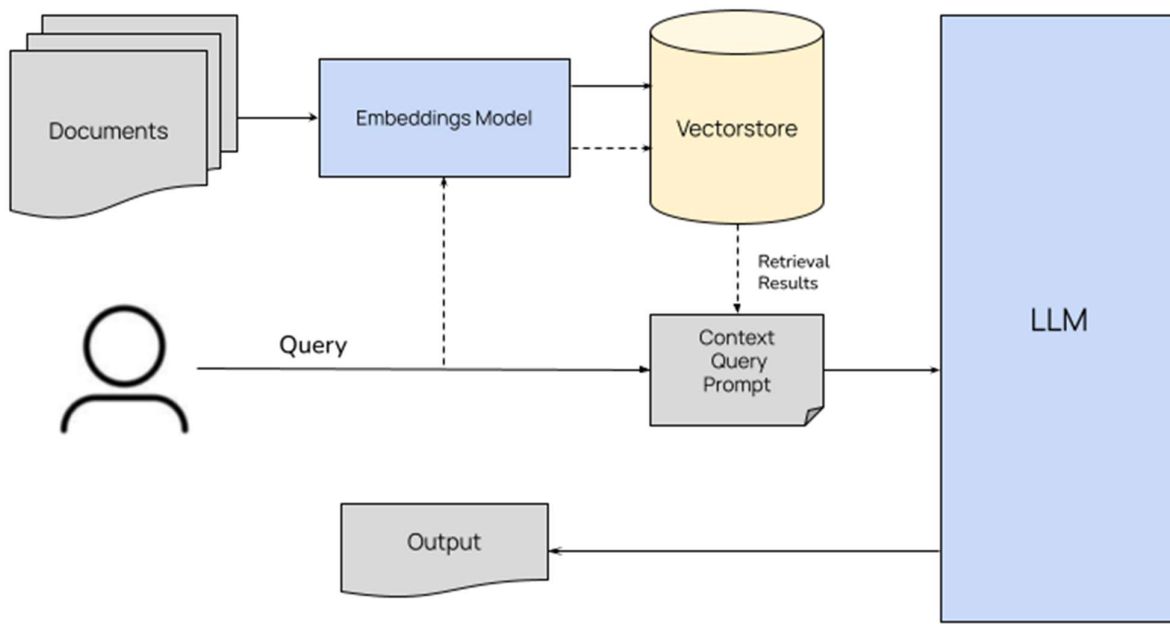
- Shyam Sunder Balaji
- Akshit Shetty
- Naveen Kumar Rudrappa Onkarappa

Objectives:

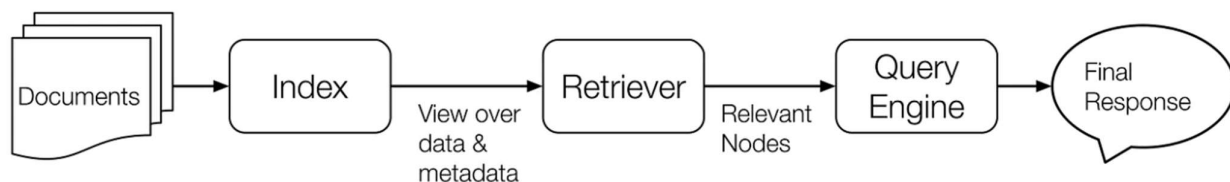
- ❖ To build a robust generative search system capable of effectively search through seven of the HDFC insurance policy documents and accurately retrieve information to answer questions from various insurance policy documents.

RAG - System Architecture:

- We employed LlamaIndex as it allows us to read files of various formats with ease to create a vector database.
- SimpleDirectoryReader and PDFReader were utilized to read the insurance PDF documents.
- The documents were parsed using SimpleNodeParser and nodes were created out of them.
- An index was created using the VectorStoreIndex of LlamaIndex.
- Finally, a query engine was constructed using the constructed index which could respond well to the user queries.
- These components were then connected by defining a query_response and initialize_conv function which allows the user to ask the model questions in chat-type setting about the insurance policies. The user is also asked to rate each response of the model.
- A pandas dataframe was constructed which included the following columns: 'Question', 'Response', 'Page', and 'Review'.



Basic RAG Pipeline in LlamaIndex:



Challenges:

- Asking complex questions sometimes resulted in not upto mark answer output. For this reason, we utilized **Sub Query Question Engine**, which breaks down the prompts into sub questions (prompts). Example below:

```

In [ ]: response = query_engine_1.query(
        "What are the key terms and conditions in the HDFC policy documents?"
    )

Generated 2 sub questions.
[document_data] Q: What are the key terms in the HDFC policy documents?
[document_data] Q: What are the conditions specified in the HDFC policy documents?
[document_data] A: Key terms in the HDFC policy documents include Policy Provisions, terms and conditions, Jurisdiction, Notice
s, Amendments, Regulations, Laws, IRDAI approval, address change procedure, Anti-money Laundering/Know Your Customer norms, and
the requirement for written communication for notices, directions, or instructions.
[document_data] A: The conditions specified in the HDFC policy documents include the right to change policy provisions/terms an
d conditions in accordance with regulations, obtaining approval from IRDAI for material changes, jurisdiction governed by India
n laws, requirements for notices to be in writing and delivered through specified means, the need for communication of address
changes by the policyholder, documents required for claim processing, issuance of duplicate policy with specific requirements,
age admission process, eligibility criteria for scheme participation, commencement of insurance based on eligibility and premiu
m payment, provision of information by the master policyholder, and the obligation to inform the insurer of any changes in the
scheme member's details within a specified timeframe.
  
```

Lessons Learnt:

- **Integration Complexity:**
 - Integrating multiple components such as LlamaIndex, Microsoft's Phi-3 model, HuggingFace models, and OpenAI's GPT-3.5 proved to be more complex than anticipated. Importance of reading the documentations and realising that each component had its own set of dependencies and configuration requirements, necessitating careful orchestration was quite beneficial to know.
- **User Feedback Integration:**
 - Incorporating user feedback mechanisms allowed for iterative improvement of the system. By analyzing the feedback, we could identify common issues and adjust the models and indexing strategies accordingly.
 - Proactive collection and analysis of user interactions can guide future enhancements and training data selection.
- **Documentation and Knowledge Transfer:**
 - Comprehensive documentation of the system architecture, integration steps, and operational procedures was crucial for knowledge transfer and future maintenance.
 - Clear and detailed documentation facilitated onboarding new team members and ensuring continuity.
- **Future Directions and Improvements:**
 - The project highlighted several areas for future improvement, including more advanced natural language understanding capabilities and enhanced user personalization to ensure the application remains effective and relevant as user needs and policy documents evolve.

END OF DOCUMENT