Insurance Agent – RAG on Insurance Data using LlamaIndex

Objective

Insurance documents are often lengthy and complex. The objective of this project is to develop an efficient Generative Search system for the insurance domain. The system will be robust and capable of accurately answering questions derived from policy documents. This will enable users to easily retrieve relevant information from numerous documents, each with a large number of pages. For this project, we will utilize a collection of sample HDFC insurance policy documents provided in previous modules.

Design Choice

We will develop a Retrieval Augmented Generation (RAG) based Agent to assist the user in their search. A RAG system can be implemented in several ways, such as using existing libraries with a custom approach or utilizing frameworks like LangChain or LlamaIndex that simplify the process of building RAG systems.

The current implementation utilizes LlamaIndex due to its emphasis on search and retrieval. Additionally, LlamaIndex is designed for efficiency, particularly in managing large data volumes. It provides a simple interface for interacting with LLMs, which facilitates the processes of searching and retrieving data.

System Design

The following section describes the system design of this application. It outlines the different layers involved in building this system and the LlamaIndex components used in each layer.

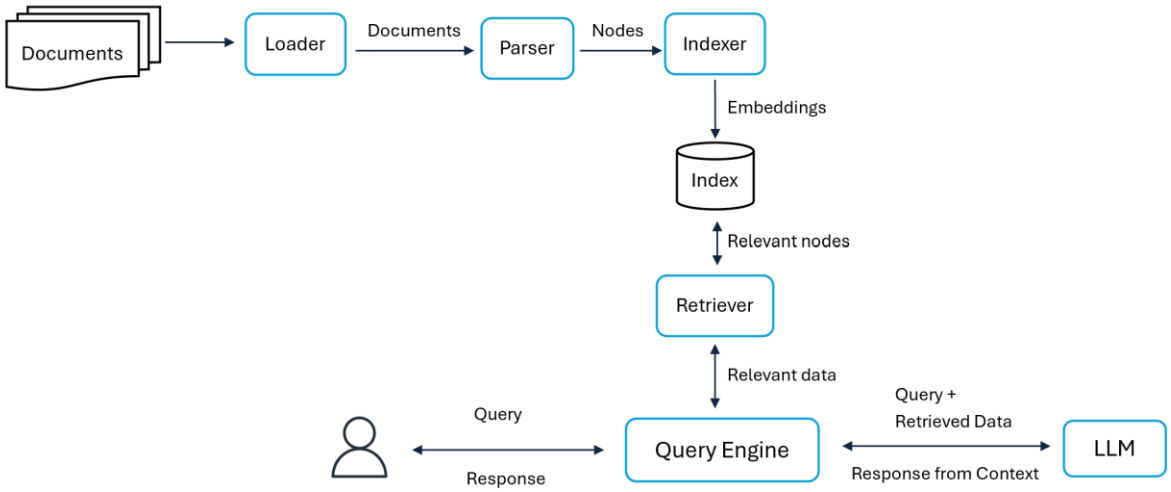

Figure 1. RAG System Using LlamaIndex

a) **Data Loading**

Data connectors and data loaders facilitate connections to various data sources and support loading of diverse file types. Currently, our policy documents are available in PDF format. For the current implementation, we utilize SimpleDirectoryReader to read multiple files from a directory. This method is employed to load files into LlamaIndex as documents for subsequent processing.

b) **Index Creation**

Each document is segmented into smaller units called nodes using the SimpleNodeParser. The indexing process involves generating nodes and linking embeddings with the data within each node. The vector storage contains the embedding vectors of processed document segments. In the current implementation, VectorStoreIndex from LlamaIndex is utilized. This approach employs LlamaIndex's default use of text-embedding-ada-002 from OpenAI. Furthermore, VectorStoreIndex relies on an in-memory SimpleVectorStore.

c) **Building Query Engine**

The Retriever is an essential component within LlamaIndex that is responsible for obtaining the most pertinent data nodes based on a user's query. The Response Synthesizer processes the user's query in conjunction with the retrieved nodes to generate a suitable response. The Query Engine serves as the central mechanism that integrates all other components, encompassing the index, retriever, and response synthesizer. It interprets the user's query, initiates the retrieval process, and produces a response that is subsequently presented to the user. The Query Engine is derived from the VectorStoreIndex established during the index creation phase. By default, LlamaIndex utilizes OpenAI's GPT-3.5-turbo model.

Steps of Project Execution:

The pre-requisites for executing this project are Python and OpenAI API key.

To get started with the Insurance Agent, execute the python notebook submitted for this project. And follow the input prompts generated by the conversation function. To end the conversation at any point type “exit” in the input prompt.

Output

The following are the screenshots of results from the system. For each query, it contains the responses from the RAG system and the corresponding citations as well from the documents provided as data store.

In [34]:
```
def testing_pipeline(questions):
  for i in questions:
    print("\nQuestion::",i)
    print("Agent Response:\n",query_response(i))
```

In [35]:
```
testing_pipeline(questions)
```

```
Question:: What is the grace period for payment in poorna suraksha policy?
Agent Response:
 The grace period for payment in the Poorna Suraksha policy is 30 days for non-monthly premium paying mode and 15 days for th
e monthly mode.
 Check further at HDFC-Life-Group-Poorna-Suraksha-101N137V02-Policy-Document.pdf page nos 9,10

Question:: In case of health insurance policy, explain the disability benefits
Agent Response:
 The disability benefits in the health insurance policy include coverage for specific critical illnesses and surgeries as lis
ted in the policy document. If the insured is diagnosed with any of the covered critical illnesses during the policy term, th
e benefit payable will be the Sum Assured, and the policy will terminate. Additionally, under the Accelerated Critical Illnes
s Option, if the insured is diagnosed with a critical illness or undergoes any of the covered surgeries listed in the policy
during the coverage term, the Accelerated Critical Illness Benefit will be payable, after which the coverage will cease and a
ll benefits will expire. The covered critical illnesses include conditions such as myocardial infarction, cancer of specified
severity, stroke resulting in permanent symptoms, Alzheimer's Disease, and others as defined in the policy document.
 Check further at HDFC-Life-Group-Poorna-Suraksha-101N137V02-Policy-Document.pdf page nos 7,14

Question:: What are the premium payment options for smart pension plan?
Agent Response:
 The premium payment options for the Smart Pension Plan include monthly premium payment mode and top-up premiums in addition
to the basic premiums within the policy term.
 Check further at HDFC-Life-Smart-Pension-Plan-Policy-Document-Online.pdf page nos 10,25

Question:: In Sanchay plus policy, what are the loan terms and conditions?
Agent Response:
 Policy loans in the Sanchay Plus policy are subject to terms and conditions specified by the company. The loan amount can be
up to a maximum of 80% of the surrender value. The current interest rate on the loan is 9.5% per annum, calculated based on t
he Average Annualised 10-year benchmark G-Sec Yield over the last 6 months, rounded up to the nearest 50 basis points, plus
2%. The interest rate is reviewed semi-annually, with changes effective from 25th February and 25th August each year. Before
any benefits are paid out, the outstanding loan amount along with the interest will be deducted, and the remaining balance wi
ll be payable. Additionally, an in-force or fully Paid-up policy cannot be foreclosed for non-repayment of the loan.
 Check further at HDFC-Life-Sanchay-Plus-Life-Long-Income-Option-101N134V19-Policy-Document.pdf page nos 11,12
```