1. **Project Background:**

We will be building a **HelpMate AI** project related to the insurance industry. An AI-powered project to extract information from insurance policy documents can revolutionize how policies are analyzed and understood. Using natural language processing (NLP), this system can efficiently scan, interpret, and extract key details such as coverage limits, exclusions, claim procedures, and premium details from the policy document

1. **Project Goal:**

To address this, we will develop **HelpMate AI**, a question-answering system that combines the power of Embedding models, RAG & Large language models to ensure accurate and reliable information delivery

1. **Problem Statement:**

Given Principal-Sample-Life-Insurance-Policy document containing information about the Group Member Life Insurance (Eligibility, Terminations, Reinstatement, BENEFITS, Member Accidental Death and Dismemberment Insurance, Dependent Life Insurance, Claim Procedures, cuisine type, area etc.), I will build a chatbot that parses the document and provides accurate answers based on the user queries

1. **Approach:**
2. **Embedding Layer:** The embedding layer generates embeddings for our text corpus and allows the RAG model to understand the meaning of the query and to generate a relevant and informative response. This is essential for a variety of tasks, such as question answering system like what we are developing
3. **Search and Rank Layer:** It ensures that the retrieved text is accurate, relevant and contextually appropriate. The search component will use a technique called semantic similarity. It is a measure of how similar two pieces of text are in terms of their meaning. The search component uses semantic similarity to retrieve documents from a knowledge base that are relevant to the user's query. The re-rank component of the search typically uses a variety of techniques to re-rank the retrieved documents
4. **Generation Layer:** The generation layer will allow the model to generate new text in response to a user's query. The generative model takes the retrieved information, synthesises all the data and shapes it into a coherent and contextually appropriate response
5. **Design choice:**

Please refer to the document attached, ‘**SystemDesign\_Sahil\_Avasthi.pptx’**

1. **Challenges faced:** Segmentation of the document and creation of a coherent response