**Text summarization:**

Design and implement a text summarization system that utilizes Natural Language Processing (NLP) techniques. The goal is to develop a model capable of summarizing lengthy documents or articles, providing a concise representation of the main ideas and key information. The assignment involves both extractive and abstractive summarization methods, allowing students to explore and compare different approaches.

Tasks:

**Data Collection and Preprocessing:**

Use the documents shared for training and evaluation.

Preprocess the data by tokenizing, removing stop words, and handling any necessary cleaning steps.

**Feature Extraction:**

Implement word embeddings for representing words or phrases.

Explore the use of named entity recognition (NER) and part-of-speech (POS) tagging for feature extraction.

**Extractive Summarization:**

Implement an extractive summarization method using techniques such as sentence scoring based on importance.

Evaluate and compare different scoring approaches.

**Abstractive Summarization:**

Explore abstractive summarization techniques using deep learning models like transformers.

Implement and compare the performance of different abstractive methods.

**Evaluation:**

Develop evaluation metrics to measure the effectiveness of the summarization model.

Conduct thorough experiments to assess the model's performance on diverse datasets.

**Comparison and Analysis:**

Compare the results of extractive and abstractive summarization methods.

Analyze the strengths and weaknesses of different approaches.

**Documentation:**

Prepare a report documenting the methodology, implementation details, and results.

Create a presentation summarizing the key findings and insights.

Here are the documents to be used for training.



**Problem Statement:**

Your task is to design and implement a chatbot that can comprehend a given document and answer questions based on its content. The chatbot should take natural language questions from users and provide relevant answers by extracting information from the document.

Tasks:

**Dataset Selection:**

Use the above documents.

Some of the questions to answer

What is a sensor fusion?

What is BMS?

What are the types of BMS?

What are the Sensor fusion uses cases and Sensor Requirements?

Document Processing:

Implement a method to preprocess and tokenize the document text, preparing it for analysis.

Question Processing:

Develop a mechanism to process and understand natural language questions from users.

Information Extraction:

Use techniques such as Named Entity Recognition (NER) or keyword extraction to identify key entities or concepts in the document.

Model Implementation:

Design and implement a chatbot model that can match user questions with relevant information in the document.

User Interaction:

Enable the chatbot to interact with users, accepting questions in natural language and providing coherent answers. don’t create any screens, just need a python code.

Deliverables:

* approach to your solution,
* able to clearly explain your thought process. (during interview)
* your python code. Share it in .py format.
* For chat bot, accuracy of answers.