ABOUT THE PROJECT

1. Construction of Datasets:

Building the training and assessment dataset is an essential part of creating a chatbot. The dataset for this RAG-based chatbot is mostly made up of questions and the answers to them that are derived from the PDF document's text. The dataset was created as follows:

The questions were chosen to address a range of subjects, definitions, and important ideas found in the PDF booklet.

Answer Generation: The information in the PDF document served as the basis for the answers. Each response offers a succinct and precise answer to the related query.
Manual Verification: Every question-answer pair was manually checked against the PDF document's content to guarantee the dataset's accuracy.

2. Selecting Assessment Measures:

Choosing the right assessment criteria is crucial to evaluating the chatbot's effectiveness. The following evaluation metrics were selected for this project:

Accuracy: Accuracy is the percentage of right answers the chatbot gives in comparison to the test dataset's ground truth answers. It gives a general idea of how well the chatbot is performing.

Precision: The percentage of accurately predicted responses among all the responses the chatbot provides is measured by precision. It aids in assessing the chatbot's capacity to deliver precise and pertinent responses.

Recall: The percentage of accurately anticipated responses among all ground truth responses in the test dataset is measured by recall. It evaluates how well the chatbot can extract pertinent data from the PDF page.

F1-score: The harmonic mean of recall and precision is the F1-score. It offers a fair assessment of the chatbot's effectiveness by taking recall and precision into account.

These metrics were chosen because they provide a comprehensive evaluation of the chatbot's accuracy, relevance, and completeness in responding to user queries based on the PDF document.

Mridul Nagpal

3. Techniques to Boost Precision:

A number of tactics are used to improve the chatbot's performance with the goal of increasing its accuracy. The following strategies were looked at:

Preprocessing: The text data collected from the PDF document was subjected to preprocessing procedures, including tokenization, lemmatization, and lowercasing.

This enhanced the efficacy of the chatbot's response creation and helped standardize the content.
Text data was transformed into numerical vectors using TF-IDF (Term Frequency-Inverse Document Frequency) vectorization. By using this method, words are given weights according to their importance and frequency in the document, which makes it easier for the chatbot to find pertinent information.

Cosine Similarity: The similarity between user queries and sentences in the PDF document was measured using cosine similarity. The chatbot determines which answer to the user's question is more pertinent by comparing the cosine similarity scores.

We hope to improve user happiness and retrieval accuracy from the PDF document by putting these tactics into practice and continuously improving the chatbot's algorithms.

In summary:

When creating a RAG-based chatbot, evaluation measures, accuracy improvement techniques, and dataset generation must all be carefully considered. Through deliberate question selection, thoughtful assessment metric selection, and efficient answer generation strategies, we may improve the chatbot's functionality and give consumers precise and pertinent information based on the PDF document's content.

Mridul Nagpal