Proposal for SHA FAQ Chatbot for Kenya

Business Understanding

This proposal aims to develop an FAQ chatbot for Kenya's Social Health Authority (SHA) to address questions related to social health insurance. The chatbot is intended to serve a diverse audience, including Kenyan citizens, healthcare providers, and government personnel, who need clear, accessible information about SHA processes, benefits, and policies. By providing a user-friendly, instant-response system, the chatbot seeks to enhance public understanding of SHA, thus supporting increased registration rates and improved access to healthcare services.

As SHA represents a significant reform in Kenya's healthcare system, the proposed chatbot will serve as an important tool in bridging information gaps and supporting the government in achieving a smooth transition. The project draws inspiration from successful chatbot implementations like the NHS chatbot in the UK, while specifically tailoring these methodologies to fit the Kenyan context.

Problem Statement

The proposed project aims to develop an FAQ chatbot capable of answering common questions about the Social Health Authority in Kenya. The chatbot will provide accurate, accessible information on SHA regulations, benefits, registration procedures, and other key topics, making crucial healthcare information more approachable for the general public.

Data Collection and Preparation

The primary data source for the chatbot will be a PDF document containing FAQs related to SHA, sourced from the Kenya Healthcare Federation (KHF) and the SHA website. This content was developed by the SHA Transition Committee with support from USAID. The data extraction and preparation will involve several key steps:

- Data Extraction: to extract FAQ data from the PDF.
- Data Cleaning and Preprocessing:
  - Duplicate and irrelevant entries will be removed to maintain data quality.
  - Text normalization will include lowercasing and the removal of special characters.
- Data Challenges: Addressing variations in question phrasing, ensuring accurate intent recognition, and managing ambiguous user queries are expected challenges during this phase.

Modeling Approach

We propose utilizing **Rasa**, an open-source framework, to build the conversational AI system. Key components included:

- NLU Model: A Natural Language Understanding model will be developed to classify user intents and extract relevant entities.
- Dialogue Management: A combination of rule-based and machine learning approaches will be used for dialogue management, ensuring accurate responses to user questions.

Evaluation Metrics

The success of the chatbot will be evaluated using the following metrics:

- Intent Classification Accuracy: Evaluates how accurately the chatbot understands user queries.
- Precision, Recall, and F1-score: These metrics will help assess the overall performance and reliability of the chatbot's responses.

The Minimum Viable Product (MVP) will initially handle a predefined set of SHA-related questions and provide accurate text-based responses, establishing a strong foundation for further improvements.

Deployment Strategy

The chatbot will be deployed using **Ngrok** to expose the locally running Rasa server to the internet. A simple **HTML** interface will be used to provide a browser-based chat experience. Users will be able to interact with the bot through a user-friendly, web-based interface without requiring additional software or messaging platforms.

Deployment Overview:

- **Ngrok** will be utilized to create a public URL, allowing external access to the Rasa server running locally.
- An HTML file will serve as the user interface, allowing users to interact directly with the bot through a web browser.
- This setup will ensure a straightforward and accessible deployment, ideal for testing and demonstration purposes.

Tools and Technologies

- **Programming Languages**: Python (pdfplumber, Rasa libraries)
- **Libraries and Frameworks**: Rasa for conversational AI, scikit-learn for NLP tasks, DIETClassifier, and TEDPolicy for advanced NLP and dialogue management
- **Deployment Platform**: HTML interface with Ngrok for external access
- **Version Control & Collaboration**: Git/GitHub for version control and collaboration among team members

Future Enhancements

The project aims to expand the chatbot's capabilities through future enhancements, which will include:

- **Additional Data Sources**: Incorporate more comprehensive datasets to cover broader aspects of healthcare information.
    - **Dynamic Question Handling**: Implement more sophisticated mechanisms for understanding and answering dynamic, open-ended questions.
- **Multi-Language Support**: multi-language support to increase accessibility for different user groups across Kenya.

Conclusion

The SHA FAQ chatbot aims to address key information needs for Kenyan citizens regarding social health insurance. By leveraging open-source tools like Rasa, combined with intuitive deployment using Ngrok and a web-based HTML interface, the project demonstrates the potential for conversational AI to enhance public engagement with healthcare initiatives. The chatbot will provide a practical solution to ensure citizens can easily access accurate, timely information about SHA, thus contributing to improved healthcare outcomes in Kenya