Data Science Capstone Project

Health Assistant

Statement of Work

**Introduction**

In a world increasingly dependent on digital solutions, the need for accessible and reliable healthcare information has become paramount. This project aims to develop a health assistant leveraging the power of AI to help users identify medical conditions based on their symptoms and recommend appropriate treatments and nearby healthcare facilities. The significance of this project lies in its potential to provide users with a unified, simple, and interactive interface that offers relevant and reliable health-related information.

**Objectives**

The primary objectives of this project are as follows:

Symptom-Based Medical Identification: Create a system that can accurately identify medical conditions based on user-provided symptoms and discomforts.

Treatment and Healthcare Facility Recommendations: Recommend appropriate treatments and healthcare facilities based on the identified medical conditions and the user's location.

Integration of Multiple Features: Unify multiple backend APIs into one comprehensive solution to enhance user experience and access to information.

Leveraging AI: Utilize AI technologies to ensure the system provides relevant and reliable results to users.

User-Friendly Interface: Develop a simple and interactive user interface for seamless interaction with the health assistant.

**Scope**

The scope of this project includes:

Symptom-Based Identification: The system will identify medical conditions based on user-provided symptoms and discomforts.

Treatment Recommendations: The system will recommend treatments and necessary actions based on the identified medical conditions.

Healthcare Facility Recommendations: The system will provide a list of nearby healthcare facilities using Google Maps API based on the user's location.

Chatbot Feature: Implement a chatbot feature that users can engage with for further assistance.

**The limitations of this project include:**

Language Support: Initially, the system will primarily support text-based interactions in English.

**Data Sources**

The project will utilize the RedHOT dataset, which includes a diverse range of health-related topics and natural claims collected from real health-related forums. The dataset provides granular annotations on claims, allowing for downstream processing.

**Methodologies**

The following methodologies will be employed in this project:

Text Classification: Text classification techniques will be used to categorize and analyze user-provided symptoms and discomforts.

Transfer Learning: Transfer learning models will be employed to enhance the accuracy of medical condition identification.

API Integrations: Integration of APIs such as ChatGPT and Google Maps to provide recommendations and location-based information.

Chatbot Implementation: A chatbot will be developed to facilitate interactive user engagement.

**Timeline**

The project timeline is as follows:

Week 1-2 (Project Initiation): Project kickoff, team formation, and initial setup.

Week 3-4 (Project Planning and Implementation): Data collection, Reddit API integration, and initial model development.

Week 5-6 (Project Refinement): Model optimization, user interface design, and integration of additional features.

Week 7-8 (Final Phase-Wrapup): Testing, user feedback incorporation, final refinements, and project documentation.

**Future Works**

Future enhancements and developments for this project include:

Speech-to-Text Input Feature: Adding support for speech-based input to enhance user accessibility.

Doctor Recommendation: Create a separate database containing doctors' details to recommend specific healthcare professionals based on user input.

Real-Time Data Updates: Hosting the dataset on Firebase and implementing automated web scraping to keep it up to date, while addressing privacy and legal considerations.