Mental Illness Chatbot – Final Submission

Term: Spring 2022

Team: Mental/Physical Illness Chatbot

Project: MAPI Chatbot

Section I: Project Details

Team Members: Tusheet Goli, Tejas Pradeep, Akshay Sathiya, Pranav Khorana, Rahul Chawla, Sanket Manjesh

TA Mentor: Tushar Aggarwal

External Mentor: N/A

Project Abstract

A primary motivation behind this project is the increasing prevalence of mental health issues in recent years, in addition to ongoing physical health issues. From the National Survey on Drug Use and Health, researchers estimate the number of adults with some degree of serious psychological distress has increased by 71% from 2006 to 2017, among young adults (Rosenberg, 2019). Furthermore, due to studies at Pew Research Center, it is known that more than 70% of people aged 18-24 use Snapchat and Instagram regularly (Ortiz-Ospina, 2019).

We wish to implement a project that would be relevant to both of these trends. A chatbot system resembling a social media interface can use machine learning to predict if the user is at risk of any mental illnesses or physical illnesses.

We define mental illnesses as illnesses mainly pertaining to mental health. We plan to support anxiety, depression, and bipolar disorder in our application. We define physical illnesses as illnesses mainly pertaining to physical health. We plan to support 41 physical illnesses (specified in the Disease Symptom Classification dataset (Patil, 2020)) in our application.

In terms of features, our application intends to provide:

- Interactive UI for a mobile application that takes in user input and visually returns feedback

- Machine learning/natural language processing techniques on the backend to analyze user symptoms
- A database connection to store user information and messages
- Accurate predictions pertaining to potential mental and physical illnesses

Research

Some of the major preliminary things we need for this project are datasets, reputable and relevant research papers and articles, and a software architecture plan. We have done extensive research and have found all the datasets and papers that we plan on using and can help us in our approach. We also have a robust software architecture plan and data pipeline for our project which has been explained in our technologies and architecture diagrams. We have been able to find some good research papers that enable us to perform novel NLP applications like sentiment analysis, and other speech/text recognition patterns to identify and classify mental illness.

Tools and Technology

Through the project, we plan on using various different tools and techniques to achieve our goals.

- Mobile App
  - Mobile applications frontend shall be built in React Native to enable us to easily use the app on both Android and iOS
  - Further data for the mobile app shall be stored on SQL, using a hosting platform such as PostgreSQL with Heroku.
  - The mobile app shall also have a backend developed on Python Flask to enable Restful API development and ease of integration with other aspects of the app.
- Server
  - We plan on using a server hosted on the cloud through free-to-use services like Heroku, with data being stored in SQL.
  - The backend code for the mobile app shall be built in Python Flask and shall serve as the link between the frontend and the server.
- Machine Learning Models
  - The machine learning models shall also be built with Python, specifically using NumPy and pandas for data cleaning and data analysis and scikit-learn libraries for the ML models to be used.
  - We chose to use Python for both the backend and the ML models for better integration between the two segments.

Data Sources

- Disease Symptom Classification (Patil, 2020)

    - The Kaggle dataset lists multiple symptoms and precautions to be taken along with weighted importance values for 41 unique diseases, including GERD, AIDS, diabetes, and gastroenteritis. We plan to train an ML model on this data to recognize these diseases from the user's descriptions of symptoms and predict if they are at risk of certain physical illnesses.

    - Link - https://www.kaggle.com/datasets/itachi9604/disease-symptom-description-dataset?select=dataset.csv

- Twitter Emotion Analysis (Merin S, 2020)

    - This Kaggle dataset contains several tweets labeled by emotion (happy, sad, anger, etc.). We plan to develop and train an ML model on this data to predict the sentiment of the user from their messages and use that information to predict if they are at risk of certain mental illnesses.

    - Link - https://www.kaggle.com/code/shainy/twitter-emotion-analysis/data

- Emotions dataset for NLP (Praveen, 2020)

    - This Kaggle dataset contains several sentences labeled by emotion (joy, sadness, fear, etc.). We plan to train an ML model on this data to predict the sentiment of the user from their messages and use that information to predict if they are at risk of certain mental illnesses.

    - Link - https://www.kaggle.com/datasets/praveengovi/emotions-dataset-for-nlp

System Architecture

Figure 1 below shows the architecture of the chatbot system and the flow of information between the user and the chatbot system.
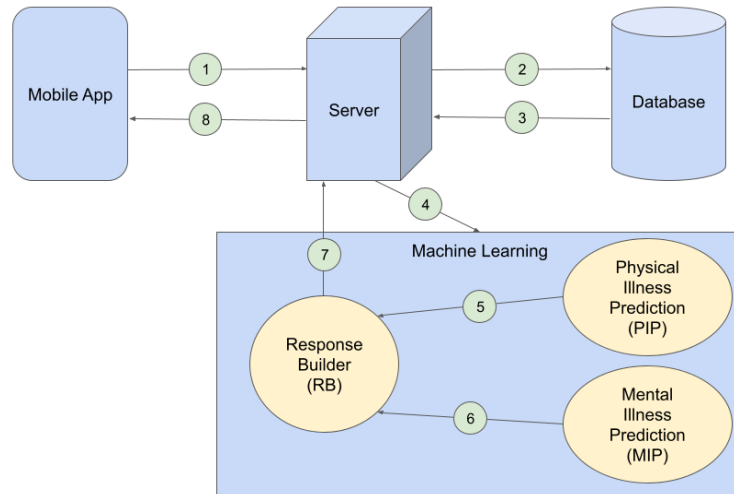
Figure 1—Architecture and information flow diagram for the chatbot system.

The chatbot system consists of a mobile app, a server, a database, and a machine learning suite. The ML suite contains an ML model for predicting physical illness (PIP), an ML model for predicting mental illness (MIP), and a response builder (RB) that builds a response to the user from the results of both models.

The information flow is described below as a series of numbered interactions that correspond to the numbered interactions shown in the diagram.

1. The user sends a message from the mobile app, which is received by the server.
2. The server stores the message in the database.
3. The server gets the last X messages from the database. The value of X will be tuned during the development and testing of the system.
4. The server sends the X messages to the ML suite.
5. The last message is passed to the PIP model to predict the physical illnesses the user may be at risk.
6. All X messages are passed to the MIP model to predict the mental illnesses the user may be at risk.
7. The RB builds a response to the user from the results from the PIP and MIP models.
8. The server sends the response to the user.

Team Member Roles & Responsibilities

Tusheet Goli - API Design, Database Implementation and Hosting, Backed Development

Tejas Pradeep - API Design, Backed Development, Project Integrations

Akshay Sathiya - Physical Illnesses Prediction ML Models, Project Integrations

Pranav Khorana - Frontend Design, React App

Rahul Chawla - Mental Illness Prediction ML Model

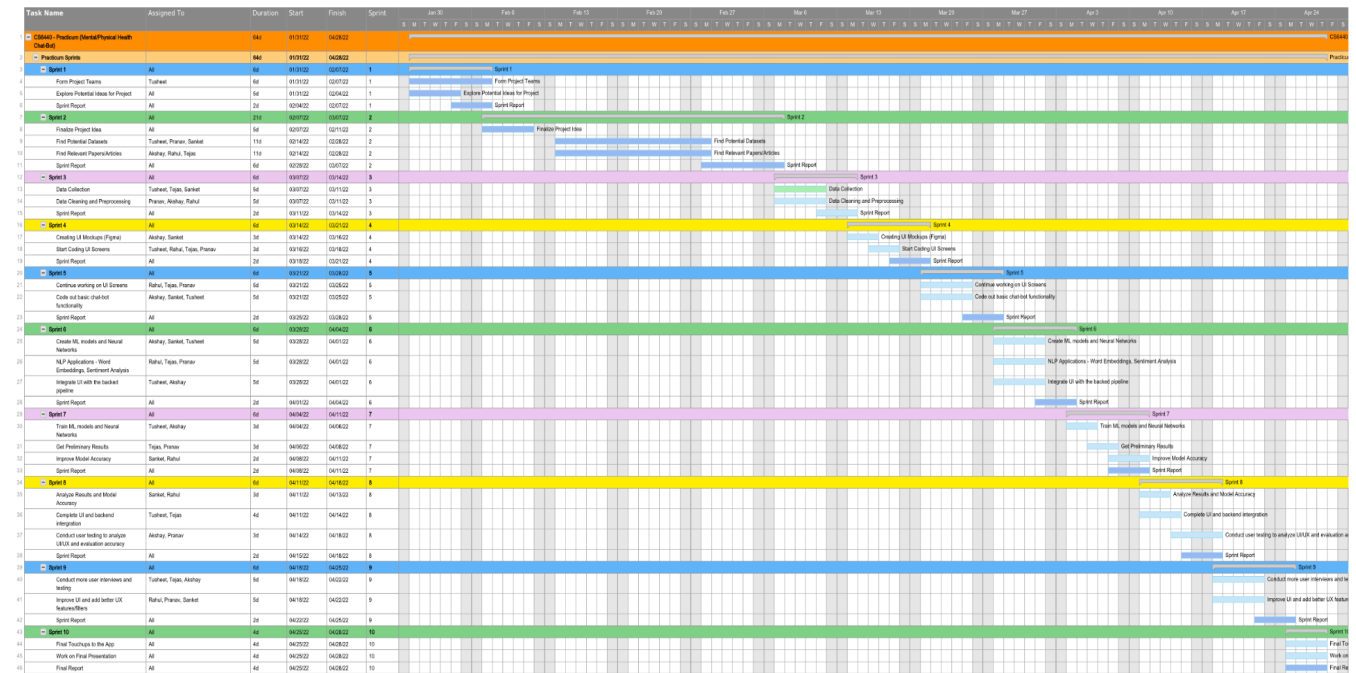Sanket Manjesh - Mental Illness Prediction ML Model

Final Gantt Chart

Figure 2—Gantt chart outlining project timeline.

Section II – Application or Solution

Final Git Commit: (cca2f78)
https://github.gatech.edu/gt-cs6440-hit-spring2022/Team-2-Mental-Illness-Chatbot/commit/cca2f78ca17cab5129c1dab385231a9bfd32f5a1

Github Link:
https://github.gatech.edu/gt-cs6440-hit-spring2022/Team-2-Mental-Illness-Chatbot

Branch: (main)
https://github.gatech.edu/gt-cs6440-hit-spring2022/Team-2-Mental-Illness-Chatbot/tree/main

Application or Solution Details

App or Solution Name: MAPI Chatbot

App or Solution URL: App is not deployed, but can be run as a regular react application from our Github repository linked below. https://github.gatech.edu/gt-cs6440-hit-spring2022/Team-2-Mental-Illness-Chatbot.

App or Solution Description

Our project idea, MAPI Chatbot, tries to provide a unique approach to self-care and self-diagnosis. We aim at diagnosing illnesses based on the symptoms that people are feeling and refer them to the right practitioners. For diagnosing mental health issues, we implemented an NLP application to analyze emotion and sentiment (Uban, 2021) in peoples' conversations with the chatbot that poses specific questions to check up on their wellbeing. Based on these evaluations, we can diagnose potential mental health issues like stress, anxiety, depression, etc. We used a similar NLP-based approach to identify physical illnesses as well. The UI of the chatbot is interactive and resembles a social media platform helping the user feel more comfortable while mentioning their symptoms, Further to facilitate natural conversation, the chatbot uses Natural Language Processing to extract symptoms from the conversation and further extract the emotions of the user messages to facilitate the mental illness predictor. Currently, the chatbot can diagnose a wide variety of physical illnesses but is limited to diagnosing three mental illnesses namely depression, anxiety, and bipolar disorder.

Section III – Project Presentation

Link to Video Presentation - https://youtu.be/B7Up9SlytSI

Link to Slides - https://docs.google.com/presentation/d/1H_-HUZiL_Ps6Vox8ky3GHBFYtxIRQo6dOesN9dASr A8/edit?usp=sharing

Section IV – Project Documentation

Final Delivery Directory: https://github.gatech.edu/gt-cs6440-hit-spring2022/Team-2-Mental-Illness-Chatbot/tree/main/Fin al%20Delivery

Final Gantt Chart:
https://github.gatech.edu/gt-cs6440-hit-spring2022/Team-2-Mental-Illness-Chatbot/blob/main/Final%20Delivery/Gantt%20Chart.png

Application or Solution Manual:
https://github.gatech.edu/gt-cs6440-hit-spring2022/Team-2-Mental-Illness-Chatbot/tree/main/Final%20Delivery/Application%20Manual

Special Instructions (if not included in the manual): None

Research Directory:
https://github.gatech.edu/gt-cs6440-hit-spring2022/Team-2-Mental-Illness-Chatbot/tree/main/Final%20Delivery/Research

Documentation Directory:
https://github.gatech.edu/gt-cs6440-hit-spring2022/Team-2-Mental-Illness-Chatbot/tree/main/Final%20Delivery/Documentation