

Assignment: Sprint 1 – Practicum Topic Exploration

Name: Tusheet Sidharth Goli

GT Username: tgoli3; GTID: 903411027; Class: CS6440-A Spring 2022

1. Problem (Topic)

Here are my top 10 project ideas that are a mix of my own, my team, and mentor project ideas.

1) Digital Therapist: A chatbot to diagnose physical and mental diseases

Source: Own/team idea. This is our top project idea that 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. This is a relevant dataset we found: [Disease Symptom Classification](#). For diagnosing mental health issues, we plan on implementing NLP applications 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. This is a relevant dataset we found: [Mental Health Chatbot](#).

2) Use ECG data to predict potential health issues

Source: Own/team idea. Our motivation behind this is to create a smartwatch app (compatible with apple watch, Fitbit, etc.) that analyzes the ECG data being collected to check for potential health issues like heart arrhythmia, heart attack, sleep apnea, etc. This can have a massive impact in mitigating health issues that can be predicted from ECG data. This is a relevant dataset we found: [ECG Heartbeat Categorization](#).

3) Emergency Department Triage

Source: Mentor project. This is based on a mentor project to create an emergency triage system that assign priorities and transfer each patient to the appropriate place for treatment. But we wanted to make this a more universal system to deal with emergency situations and natural calamities. More information regarding this is in the Mentor Projects List (Fall 2021).

4) Self-Dermatology Care

Source: Own/team idea. Our motivation behind this app is to give people access to self-care and potentially screen for any skin/derma-related issues. We intend on using a trained CV model that can classify and diagnose potential skin issues based on the images that people upload. This is a relevant dataset we found: [Derma Classification](#).

5) Detecting population at risk or predicting future illnesses based on public health data

Source: Own/team idea. This is intended to be a data analysis and ML for health application where we plan on going scouring through publicly available health data like chest x-rays, MRI/CAT scans, etc. to detect population at risk for new diseases or other public health issues. These are some relevant datasets: [Brain MRI](#), [Chest X-Ray](#), [COVID-19](#).

6) Analytical application regarding multiple illnesses

Source: Own/team idea. We wanted to build an analytical application to indirectly screen patients for serious health conditions like COVID, lung disease, heart disease, etc. based on information like nutrition, smoking, vaping, flu, vaccinations, etc., and demographic data.

7) Wound Care Tracking

Source: Mentor project. This is a CV application that helps users keep a track of the stages of their healing wound and how to better care for it. This is a relevant dataset: [Wound Care](#).

8) Online Game to Diagnose Dyslexia

Source: Mentor project. We want to create an online screening tool for Dyslexia that presents images to users and detects potential dyslexia based on their selection. It is based on this scientific article that we read about: [Dyslexia Screening Tool](#).

9) Nutrition Tracker

Source: Mentor project. This is an app to track nutritional intake to ensure people are getting the correct nutrition daily. Unlike other apps, we wanted to center this app around post-health-condition healing that asks for specific nutritional/food intake and medication.

10) Obesity Tracker

Source: Mentor project. For people who are afraid of obesity or have been diagnosed with obesity, we wanted to create an app where people can track their calorie and nutritional intake to ensure lowering their risk of obesity.

2. Area of Focus

Based on my main project idea, Digital Therapist, our area of focus is going to mainly be Machine Learning and Natural Language Processing for identifying features for diseases classification, and sentiment and emotion analysis indicators in NLP to identify potential mental health issues. We want to mainly focus on self-care and self-diagnosis and create scientific tools for the betterment of the general public health.

3. Background and Significance

Millions of people in the U.S. are affected by mental illnesses every year. Here are some statistics to show the seriousness of mental health issues in the U.S. (NAMI, 2021):

- 21% of U.S. adults experienced mental illness in 2020 (52.9 million people). 1 in 5 adults.

- 5.6% of U.S. adults experienced serious mental illness in 2020 (14.2 million people).
- 16.5% of U.S. youth aged 6-17 experienced a mental health disorder in 2016 (7.7 million people).

Even with such high rates of mental health issues, this is taken for granted as it comes along with a lot of stigmas. These stats show how few people receive treatment for mental health:

- 46.2% of U.S. adults with mental illness received treatment in 2020.
- 64.5% of U.S. adults with serious mental illness received treatment in 2020.
- 50.6% of U.S. youth aged 6-17 with a mental health disorder received treatment in 2016.
- The average delay between onset of mental illness symptoms and treatment is 11 years.

With such high rates of mental health disorders, there has never been a greater significance for de-stigmatizing and timely diagnosing mental health issues and self-diagnosis.

4. Proposed Solution or Idea

Understanding the significance of mental health and self-care, we intend to create an online tool not only for symptom-disease diagnosis but also to detect mental health issues. For diagnosing illnesses, people can enter the symptoms with the chatbot, and we aim at providing an accurate diagnosis of the condition that they are facing or referring them to the right practitioners. We plan on implementing ML/NLP applications to analyze emotion and sentiment in peoples' conversations with a 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. This will help de-stigmatize and raise public awareness about mental health by creating a culture of self-care and self-healing.

5. Complexity or Effort

The complexity of our implementation arises from our ML and NLP model which is the backbone of our application. Training our ML model to accurately classify/categorize diseases based on symptoms, implementing a sentiment/emotion analysis in conversations with the chatbot, coming up with the correct algorithms for analyzing patient data, and setting up the data pipeline for collection are some major technical complexities of the app. Although these might seem complex to implement, they are definitely achievable goals.

6. Tentative Team Members & Roles

We have a team of six members – Tusheet Goli, Tejas Pradeep, Sanket Manjesh, Pranav Khorana, Rahul Chawla, Akshay Sathiya. Since our project idea has a lot of components that involve ML, NLP, backend and frontend architecture, etc. we plan on splitting our team based on functionality. Since we have a mental health aspect and a symptom-disease diagnosis, we can divide our team into two sub-teams to handle each of these two aspects.

References

- AI, A. I. F. (2022, February 2). *Covid-19 open research dataset challenge (cord-19)*. Kaggle. Retrieved February 6, 2022, from <https://www.kaggle.com/allen-institute-for-ai/CORD-19-research-challenge>
- Buda, M. (2019, May 2). *Brain MRI segmentation*. Kaggle. Retrieved February 6, 2022, from <https://www.kaggle.com/mateuszbuda/lgg-mri-segmentation>
- Centers for Disease Control and Prevention. (2021, June 28). *About mental health*. Centers for Disease Control and Prevention. Retrieved February 5, 2022, from <https://www.cdc.gov/mentalhealth/learn/index.htm>
- Fazeli, S. (2018, May 31). *ECG Heartbeat Categorization Dataset*. Kaggle. Retrieved February 6, 2022, from <https://www.kaggle.com/shayanfazeli/heartbeat>
- Goel, S. (2020, June 24). *Dermnet*. Kaggle. Retrieved February 6, 2022, from <https://www.kaggle.com/shubhamgoel27/dermnet>
- Mental health by the numbers*. NAMI. (2021, March). Retrieved February 5, 2022, from <https://www.nami.org/mhstats>
- Mooney, P. (2018, March 24). *Chest X-ray images (pneumonia)*. Kaggle. Retrieved February 6, 2022, from <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>
- Patil, P. (2020, May 24). *Disease symptom prediction*. Kaggle. Retrieved February 6, 2022, from <https://www.kaggle.com/itachi9604/disease-symptom-description-dataset?select=dataset.csv>
- Prasath, N. (2020, October 2). *Mental health FAQ for Chatbot*. Kaggle. Retrieved February 6, 2022, from <https://www.kaggle.com/narendrageek/mental-health-faq-for-chatbot>
- Rauschenberger, M., Baeza-Yates, R., & Rello, L. (1AD, January 1). *A universal screening tool for dyslexia by a web-game and machine learning*. Frontiers. Retrieved February 6, 2022, from <https://www.frontiersin.org/articles/10.3389/fcomp.2021.628634/full>
- Uban, A.-S., Chulvi, B., & Rosso, P. (2021). An emotion and cognitive based analysis of mental health disorders from social media data. *Future Generation Computer Systems*, 124, 480–494. <https://doi.org/10.1016/j.future.2021.05.032>
- Taher, M. (2021, December 2). *Wound data*. Kaggle. Retrieved February 6, 2022, from <https://www.kaggle.com/mohamadtaher/wound-data>

Zunic, A., Corcoran, P., & Spasic, I. (2020, January 28). *Sentiment Analysis in health and well-being: Systematic review*. JMIR medical informatics. Retrieved February 5, 2022, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7013658/>