# CSE445: Machine Learning

Project Proposal - Disease Prediction System

# Group Members:

| Name: | ID: |
| --- | --- |
| Tasneem Mahmud | 1731893642 |
| Durdana Kamal | 1813355042 |
| Sumit Kumar Das | 1731847642 |

# Introduction:

A few years back, who would have thought that, we would be able to predict diseases using computers. Thanks to machine learning, we are able to do it now.

With the advances of machine learning we are able to predict such data which are critical and might as well save a life or correctly predict a disease for early hospital treatment.

The project that we will work on, will be predicting disease beforehand and it will make computer science and medical science to work in harmony for the better future.

# Abstract discussion

Using machine learning to predict diseases beforehand will help us create a healthcare related system. We wish to create a project that will be able analyze the datasets and predict whether the person has a particular disease or not.

We wish to implement machine learning algorithms on the dataset we have selected for our project and predict the disease. Its implementation is done using the python programming language.

We will be using columns containing diseases, their symptoms , precautions to be taken,  their weights(level of effectiveness per 2 days) and so on.

The data set we will be using, might have 42 diseases for processing. Based on the symptoms, age, and gender of an individual, the diagnosis system gives the output as the disease that the individual might be suffering from. Our diagnosis model can act as a doctor for the early diagnosis of a disease to ensure the treatment can take place on time and lives can be saved.

# Data set references

**Contents**

1. Columns of various diseases
2. Columns of multiple symptoms of each disease
3. Columns of multiple precautions/treatments of each disease
4. Descriptions and definitions of each disease
5.

We are also planning on including some numerical data within our dataset, to improve the accuracy of our prediction model.

# Data set references

Citation:


1. Patil, P. (2020, May). Disease Symptom Prediction, Version 2. Retrieved from https://www.kaggle.com/itachi9604/disease-symptom-description-dataset?select=dataset.csv

# Methodology

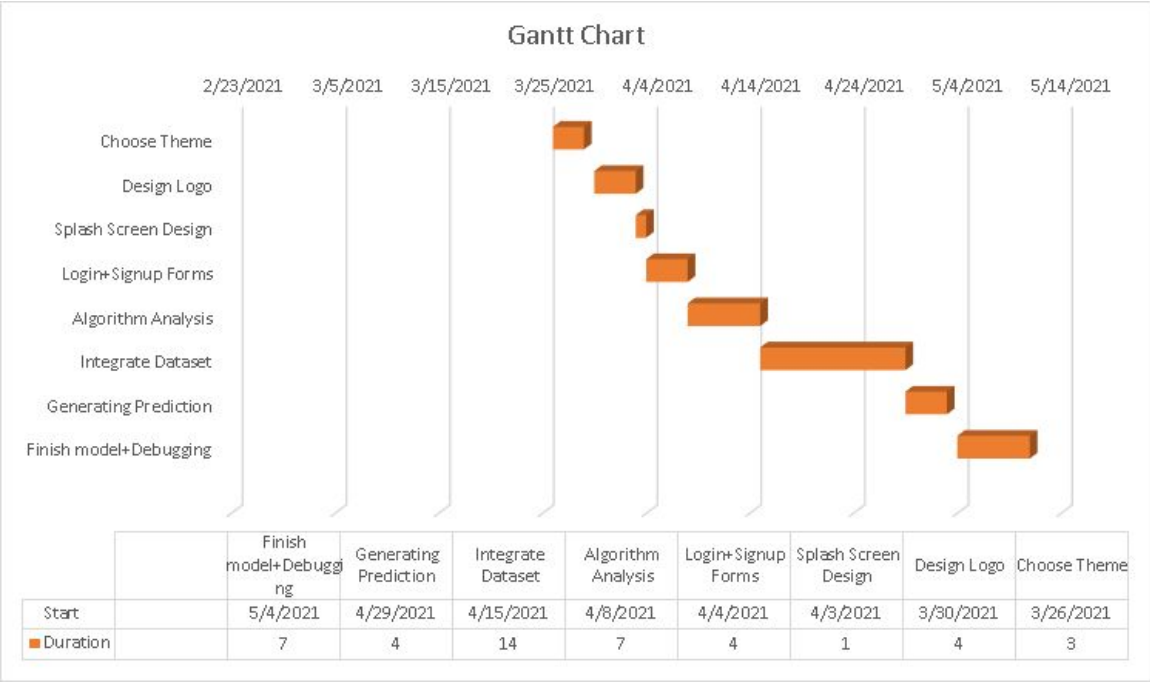
This predictive model predicts what disease the user might have. In order to do that it needs to match symptoms. This is a classification predictive model. This will be an android app. We will integrate the android app with a very well structured dataset. That dataset will probably play the most important role in the prediction process.

The dataset consists of 42 diseases. Few diseases were picked based on medical reports of patients, and then the ones people are most prone to having were picked. Diseases might present themselves in different ways in a certain set of people. But they always have a few common effects in people. The dataset consists of those common symptoms. Especially, when it comes to COVID you cannot really use all the symptoms people have shown. Adding excess information makes the prediction unreliable. Also, viruses like the covid-19 can present themselves with common flu like symptoms.

# Methodology

Our app will ask you about your symptoms. Then based on your response it checks the dataset for matching symptoms. Once the app has found a match, it will give you a diagnosis on what disease you might have. It will also show you the full list of symptoms so that if you actually have the disease, you will know what other symptoms to expect in the near future.

# Gantt Chart




| | | Finish model+Debugging | Generating Prediction | Integrate Dataset | Algorithm Analysis | Login+Signup Forms | Splash Screen Design | Design Logo | Choose Theme |
|---|---|---|---|---|---|---|---|---|---|
| Start | | 5/4/2021 | 4/29/2021 | 4/15/2021 | 4/8/2021 | 4/4/2021 | 4/3/2021 | 3/30/2021 | 3/26/2021 |
| Duration | | 7 | 4 | 14 | 7 | 4 | 1 | 4 | 3 |

# Conclusion

Illness isn’t something people take lightly. Often the panic is more harmful than the disease itself. Especially when an unknown life threatening disease is spreading worldwide so quickly that it gives us flashbacks of the 1910s, when the destructive wave of the spanish flu created havoc all over the world. This virus, widely known as the coronavirus, has already taken millions of lives and has left the rest in panic. Our goal is to help people understand the disease more, so that they can save their own lives and of those they care about, by taking necessary precautions. And in the future, no one should live their lives oblivious to the diseases out there. A timely prediction of a disease will help to slow down the disease, if not prevent it.