**Overview:**

You are given a medical dataset for which you are required to divide the dataset into multiple clusters such that you are able to analyze these clusters, explain the cluster formation and can predict the cluster if a new data point(user) input is given.

The tasks are basically divided into 3 parts:

1. Cluster the dataset.
2. Analyzing the cluster results.
3. Create a django api which can take datapoint(s) as input and returns their cluster labels(inference).

**Dataset:**

Download the train and test dataset from this [link](https://www.kaggle.com/competitions/onelasttrain/data) .

**Notes:**

1. The goal of this assignment is to understand the candidates data related thought process and problem solving skills as well as familiarity with python and machine learning.
2. Try to finish as many parts as you can, don’t worry much about the accuracy.
3. Provide your approach in a document with all the assumptions, results and how to run the application.
4. Just have fun while working on this!