LAB REPORT

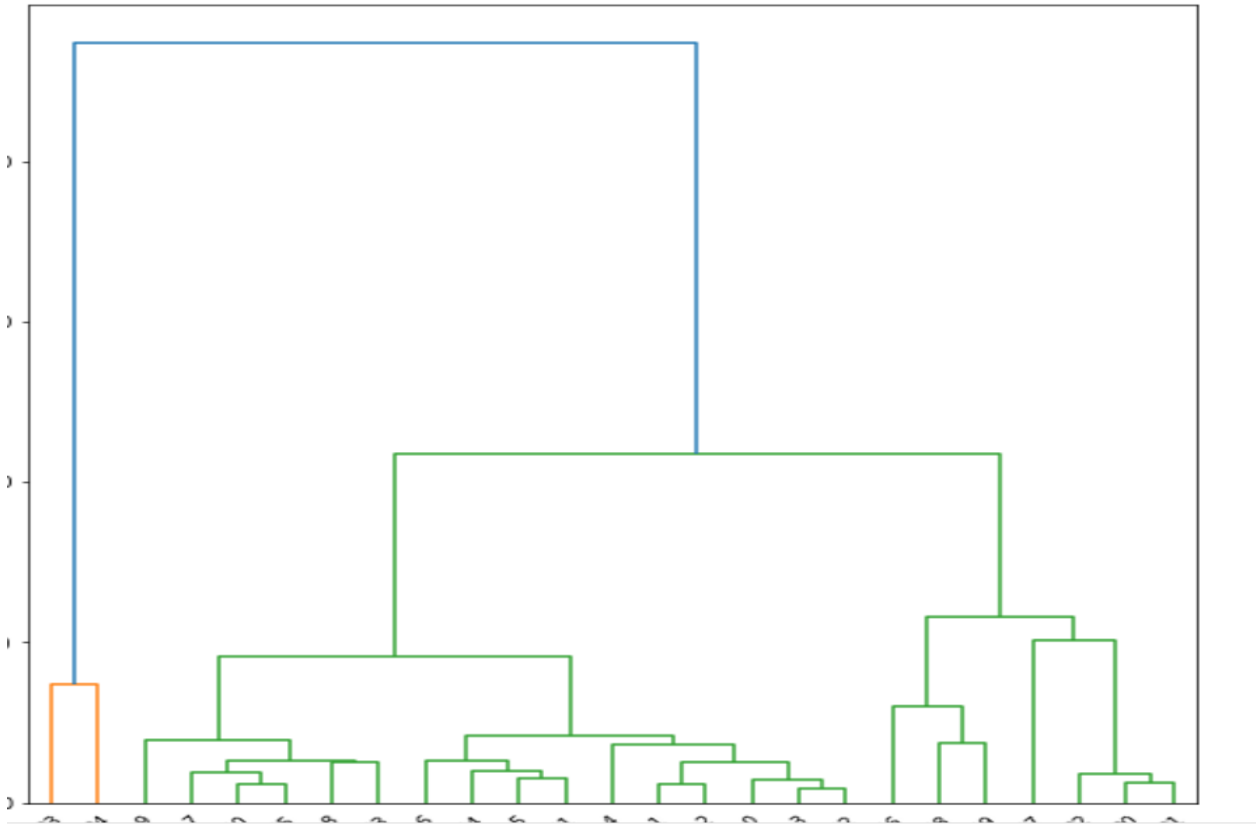
Name-Devyani Gorkar
Roll no. -B20ME027

About code:

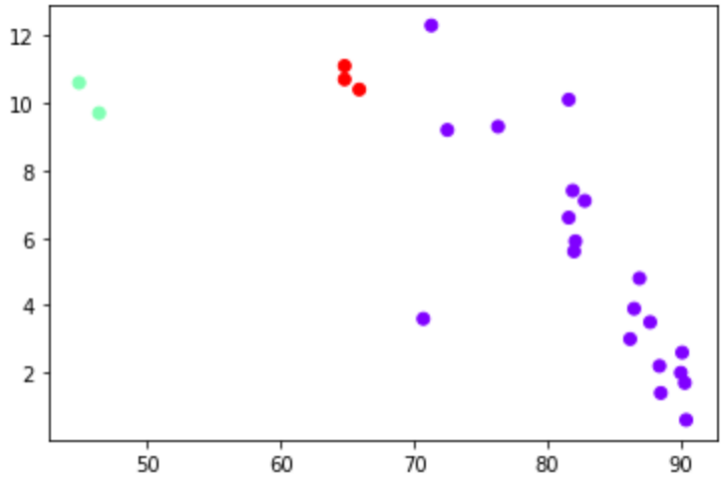
- Df is the dataframe of milk.csv.Labellist is the list of all the animals in the dataset.X numpy array of the df excluding the 1st column.
- Dendrogram is plotted using the scipy library for agglomerative clustering
- Using sklearn library clustering is done with different values of n_clusters and the graph is plotted for the same.
- While implementing k_means clustering from scratch we choose k randomly.The first loop iterates and updates the centroids until they converge.Centroid_array will be an array which will first consist of random cluster points.
- In the second loop we will iterate through the rows of X.We will find euclidean distance of each datapoint of X w.r.t every element of centroid array and then append then the minimum distance into a list l, which is emptied for each datapoint.
- We create another list named as initial_cluster which will store the index of the cluster point having minimum distance with the datapoint ,in the 1st column.In the 2nd column it will store the clusterpoint at minimum distance.3rd column will store the index of the datapoint.
- P is the numpy array of initial_cluster.m is a list of unique indexes of cluster points present in p[:,0].cl is a dictionary which will store the index of the cluster center as key and the data points of that cluster as a numpy array.
- In the last loop we can see that we update the value of centroids array by the column-wise mean of the data points.
- Y is the array of the indexes of each cluster center to which a datapoint of X belongs.SImilar to the array obtained after using predict(X) in sklearn library.
- Atlast we plot the graph of data points belonging to the cluster centers.Black coloured points are cluster centers others are clusters.

  Google colab link:
  https://colab.research.google.com/drive/14-CCoXgFWSS97K2G3gRVQc6ksUxH2nU8?usp=sharing

Dendrogram:



Agglomerative clustering :No. Of clusters =3



K means clustering:No. Of Clusters=4