

Clustering

1) Given a training dataset

X_1	X_2
0.204000	0.834000
0.222000	0.730000
0.298000	0.822000
0.450000	0.842000
0.412000	0.732000
0.298000	0.640000
0.588000	0.298000
0.554000	0.398000
0.670000	0.466000
0.834000	0.426000
0.724000	0.368000
0.790000	0.262000
0.824000	0.338000
0.136000	0.260000
0.146000	0.374000
0.258000	0.422000
0.292000	0.282000
0.478000	0.568000
0.654000	0.776000
0.786000	0.758000
0.690000	0.628000
0.736000	0.786000
0.574000	0.742000

Writing the training program using **KMeans** in **scikit-learn** library to partition this dataset into 4 clusters. The program visualises color datapoints (the color corresponds to the cluster id).

2) Writing the training program using **KMeans** in **scikit-learn** library to partition **Iris dataset** into 3 clusters. The program visualises color datapoints in the matrix of scatterplots 2D (the color corresponds to the cluster id).