Alptekin Orbay
2015400252
2 April 2017

CMPE 462 HW2

Mixture of Gaussians

I initialise the means, covariances and prior probabilities by using k_means algorithm. K_means returns as clusters, from those, initial means are the centroids, covariance matrices are the covariances of clusters and priors are the size of the clusters. To determine the convergency, I get the 2-norm of the differences of old and new means and sum them to compare a small epsilon such as 0.0001. All combinations are tested in for loop and maximum error rated one is held as optimum. Optimum combination is tested on test set and the results are illustrated below.

Cluster Sizes = Class1 : 1,Class2 : 3, Class3 : 2

Prediction Error : 0.23

Actual \ Prediction	Class 1	Class 2	Class 3
Class 1	385	105	10
Class 2	58	386	56
Class 3	7	111	382

KNN

I separated the data in three sets training set, validation set and test set. For each instance, the distances with all training set instances are calculated. For k < 10, I picked the minimum distances sequentially. Otherwise, I sorted the distances and take the first k element as log1500 > 10 for N*logN < k*N. From those k element, the mode of the classes are assigned to the class estimate for majority vote.

Best k : 40

Actual \ Prediction k=1	Class 1	Class 2	Class 3
Class 1	370	103	27
Class 2	111	278	111
Class 3	26	109	365

Prediction Error : 32.5

Actual \ Prediction k=10	Class 1	Class 2	Class 3
Class 1	393	89	16
Class 2	97	343	60
Class 3	13	100	387

Prediction Error : 25.2

Actual \ Prediction k=40	Class 1	Class 2	Class 3
Class 1	387	98	15
Class 2	87	348	65
Class 3	8	102	390

Prediction Error : 25