ASSIGNMENT 2

CS5691 Pattern Recognition and Machine Learning

CS5691 Assignment 2

Team Members:	
BE17B007	N Sowmya Manojna
PH17B010	Thakkar Riya Anandbhai
PH17B011	Chaithanya Krishna Moorthy

Indian Institute of Technology, Madras

Contents

1 Dataset 1A

1.1 K-nearest Neighbors Classifier

1.2 Naive-Bayes classifier

1.2.1 Same Covariance Matrix ($\sigma^2 I$)

1.2.2 Same Covariance Matrix (C)

1.2.3 Different Covariance Matrix

2 Dataset 1B

2.1 K-nearest Neighbors Classifier

2.2 Bayes Classifier, GMM, full covariance

2.2.1 Equations

Th initialization is done as follows for each class:

- Cluster initialization is using `kmeans` clustering.
- The relative number of points in each cluster N_q and weightage w_q for each cluster is calculated.
- The responsibility $\gamma_{n,q}$ is then calculated, followed by mean μ_q and covariance C_q is calculated.

The parameters are then updated sequentially through the:

- Expectation-step: $\gamma_{n,q}$ is updated.
- Maximization-step: μ_q, C_q, N_q and w_q are updated.

The stopping criterion used is $\Delta(\text{likelihood}) < \texttt{tol}$. The `tol` we considered is 10^{-5}.

Based on the accuracies obtained on the training, validation and test dataset, the best q_i for the three classes has been chosen as 5.

2.2.2 Training and Validation Accuracy

The training and validation accuracies obtained for varying q_i for each class is as follows:

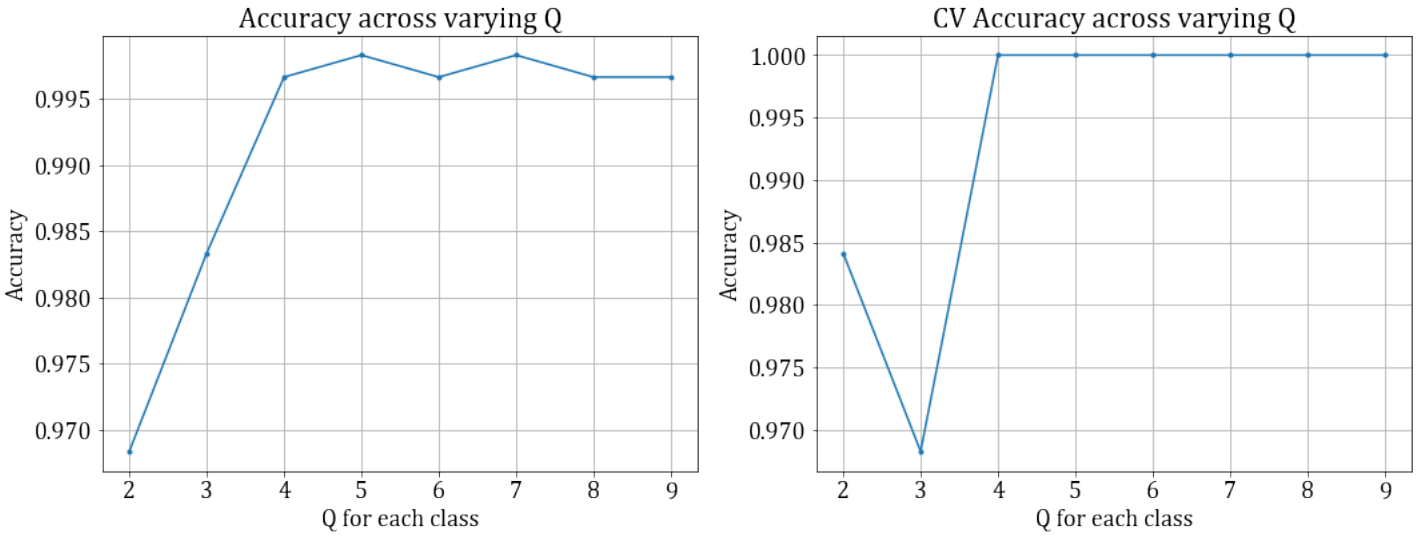

Figure 1: Training and Validation accuracy across q_i, on the left and right respectively

2.2.3 Testing Accuracy

The testing accuracy obtained for varying q_i for each class is as follows:

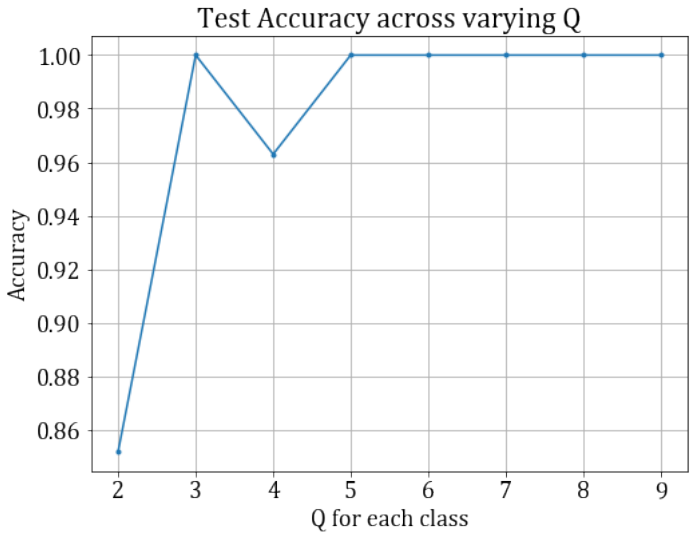

Figure 2: Testing accuracy across q_i

2.2.4 Contour Maps and Decision Surfaces

The contour maps and decision surfaces obtained, with $q_i = 5$ are as follows:

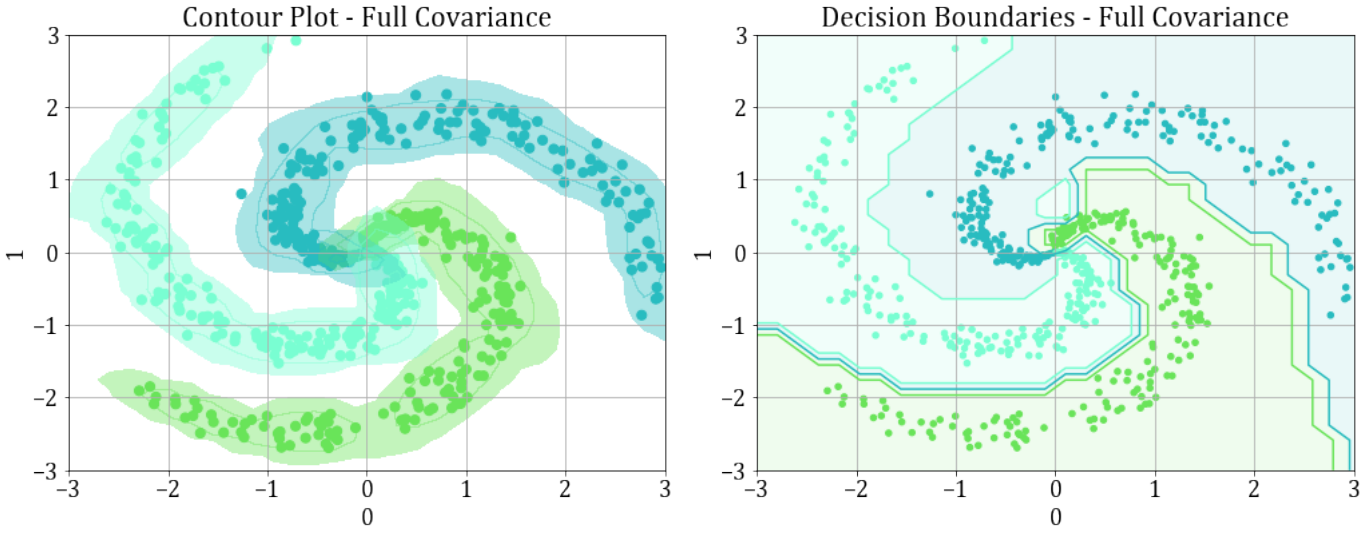

Figure 3: Contour Maps and Decision Surfaces obtained for $q_i = 5$, on the left and right respectively

2.3 Bayes Classifier, GMM, diagonal covariance

2.4 Bayes Classifier, KNN

3 Dataset 2A

3.1 Bayes Classifier, GMM, full covariance

3.2 Bayes Classifier, GMM, diagonal covariance

4 Dataset 2B

4.1 Bayes Classifier, GMM, full covariance

4.2 Bayes Classifier, GMM, diagonal covariance