

CS5691: PRML - Assignment 02

Group 20

CS24M003 Pradeep Peter Murmu , GE24Z009 Naveen Seth Hanig

September 14, 2024

Dataset-1

KNN

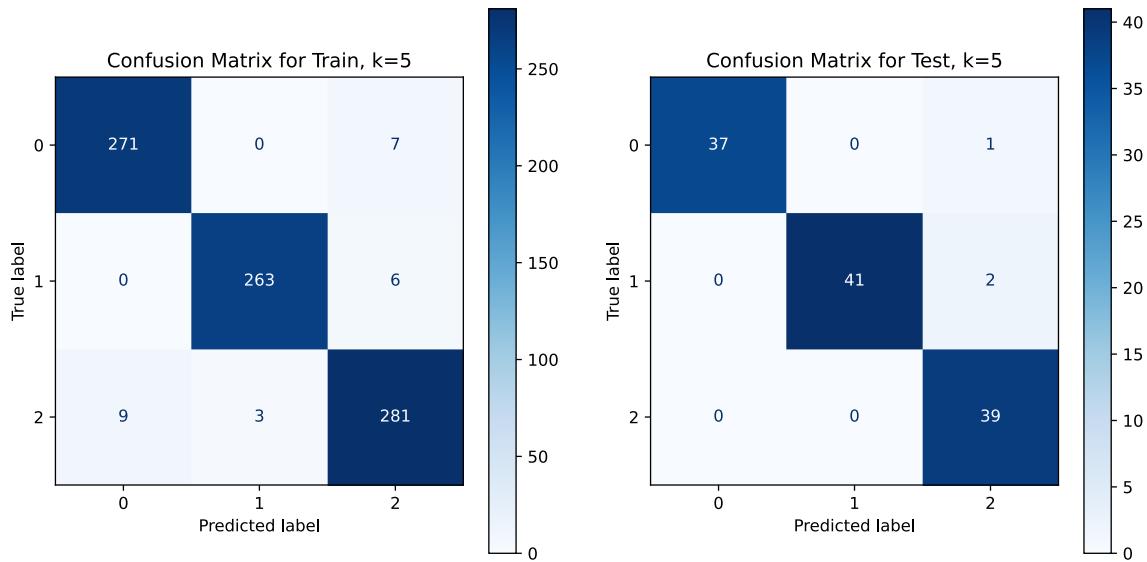
Accuracies

K	Train Accuracy	Validation Accuracy	Testing Accuracy
1	1	0.9625	0.9500
5	0.9702	0.9625	0.975
9	0.9678	0.9583	0.975

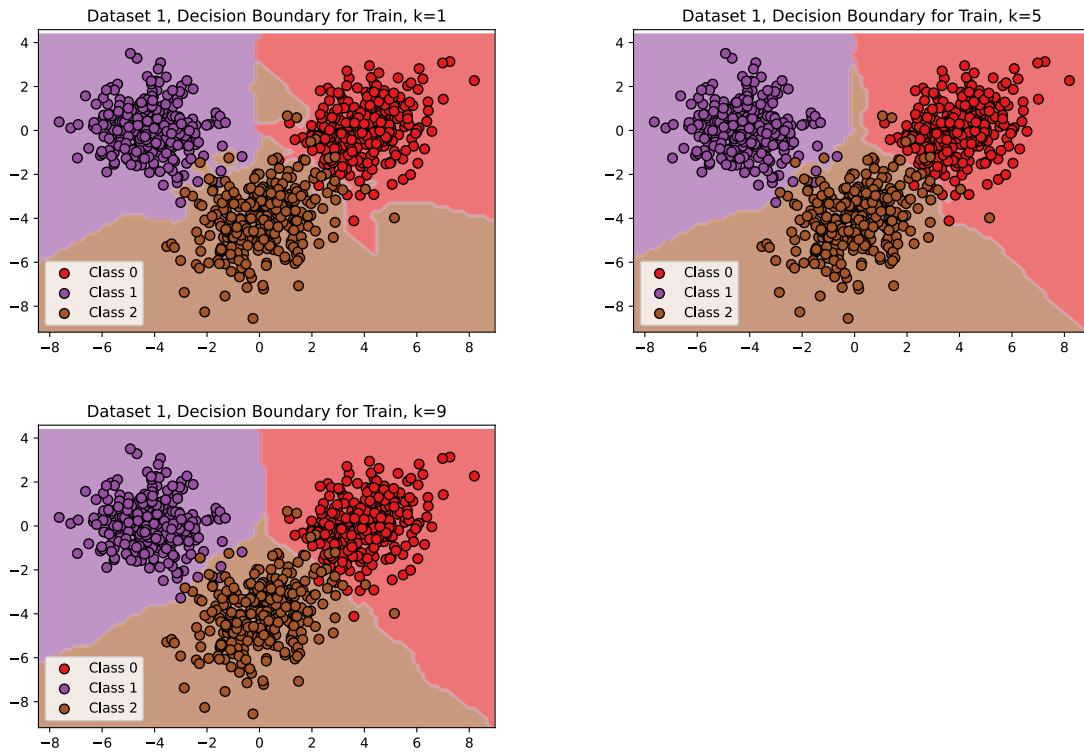
Table 1: KNN Accuracy

The validation accuracy for $k = 1$ and $k = 5$ is equal. Therefore, we use the testing accuracy as a secondary metric for determining the most accurate model. The model using $k = 5$ performs the best.

Confusion Matrices



Plots



Bayes classifier with a Gaussian distribution

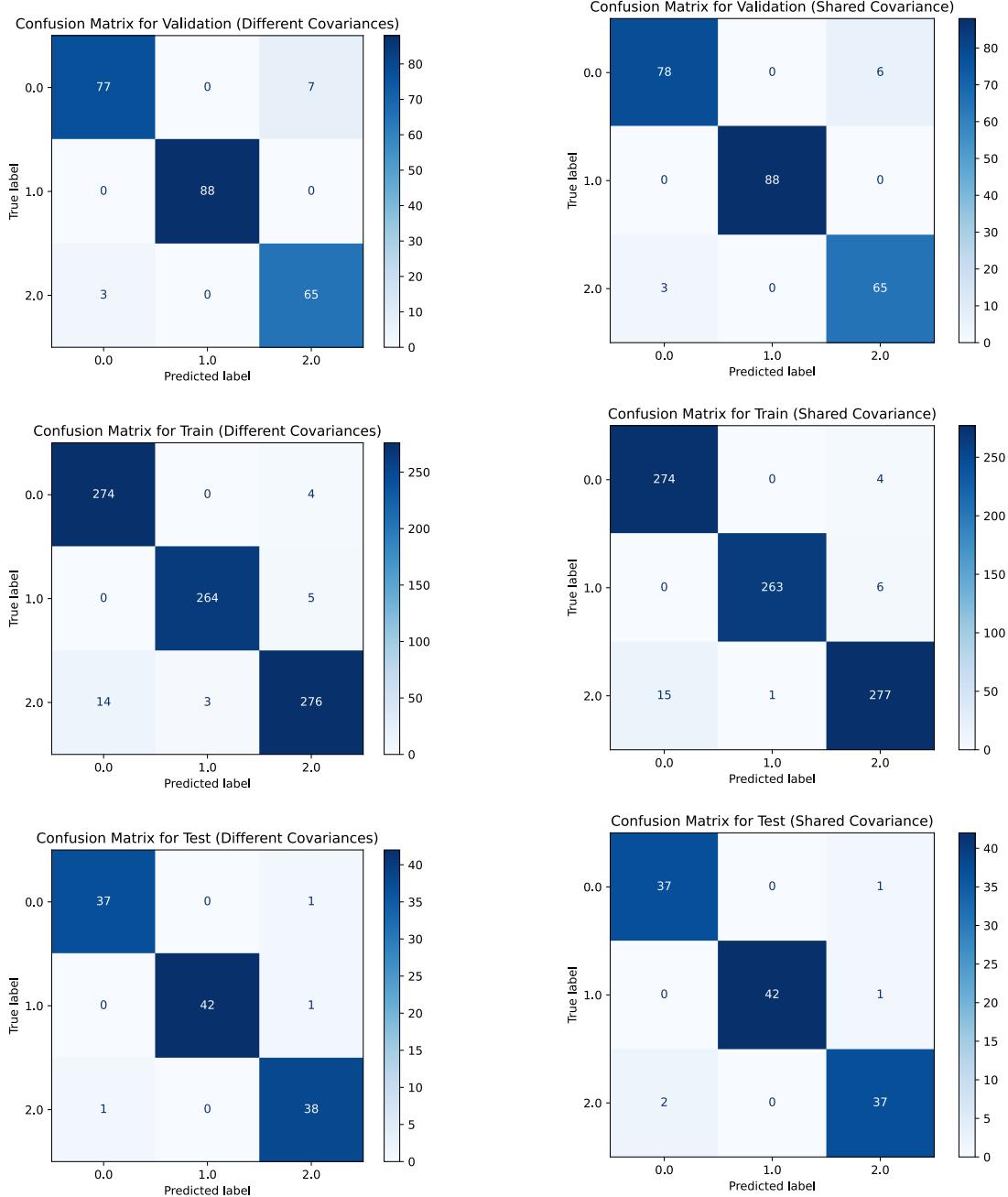
Accuracies

Type	Train Accuracy	Validation Accuracy	Testing Accuracy
Bayes Gaussian with different covariances	0.9690	0.9583	0.9750
Bayes Gaussian with shared covariance	0.9690	0.9625	0.9667

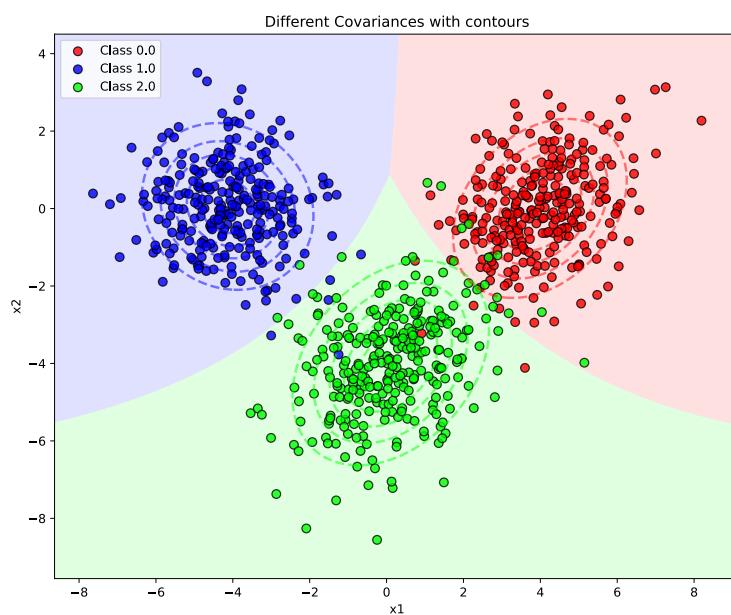
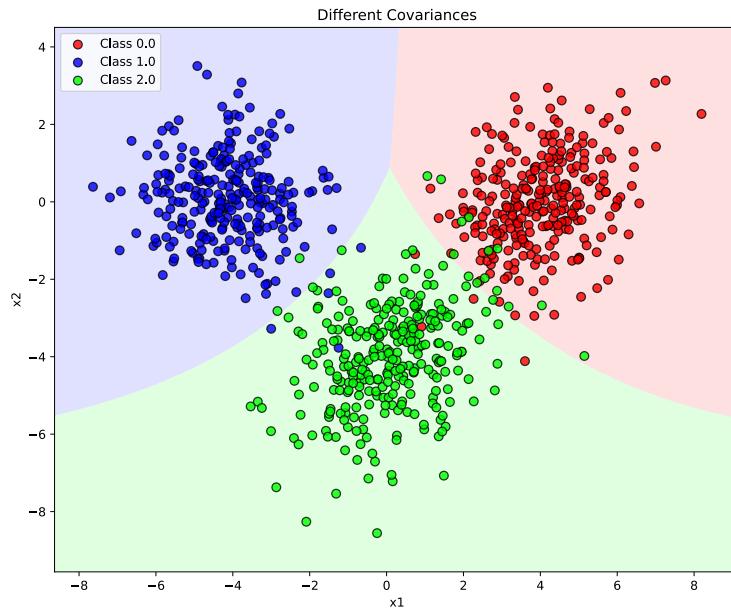
Table 2: Bayes Gaussian Accuracy

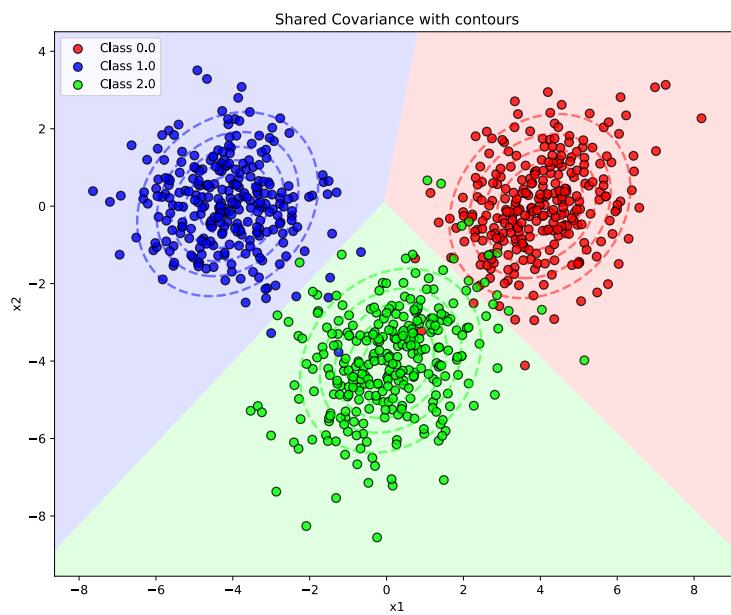
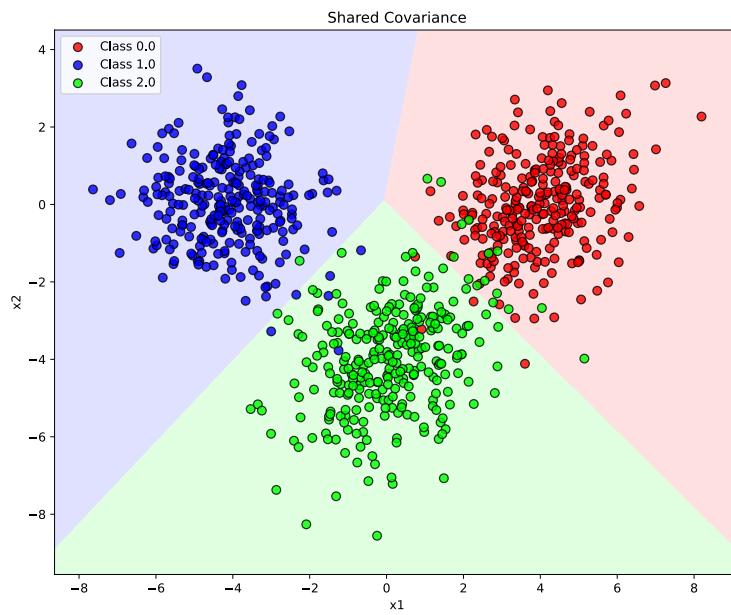
Best Model- Bayes Gaussian with shared covariance due to higher accuracy for validation data

Confusion Matrices



Plots





Dataset-2

KNN

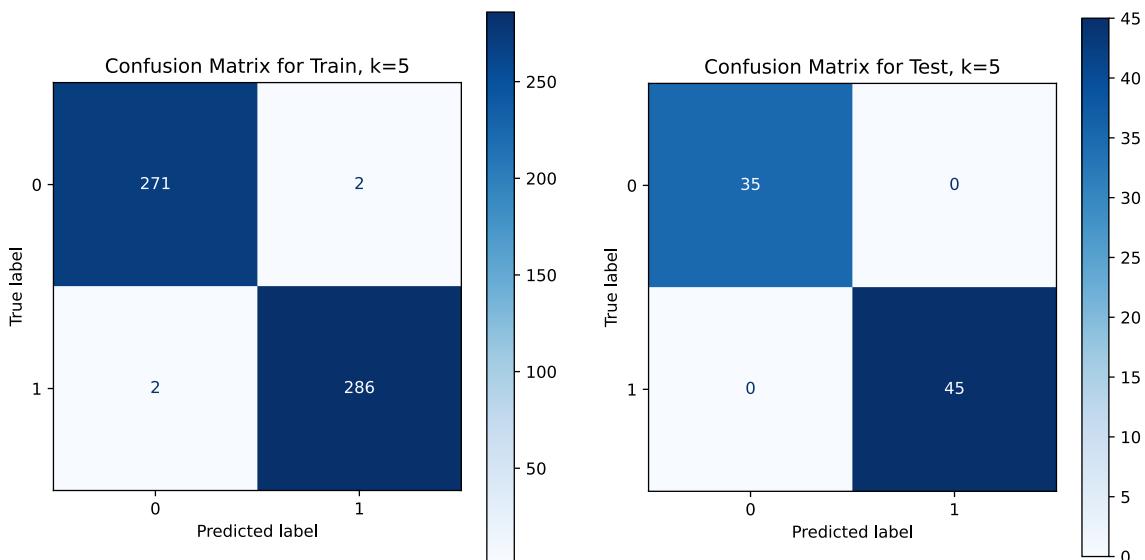
Accuracies

K	Train Accuracy	Validation Accuracy	Testing Accuracy
1	1	0.9811	1
5	0.9928	1	1
9	0.9893	0.9937	1

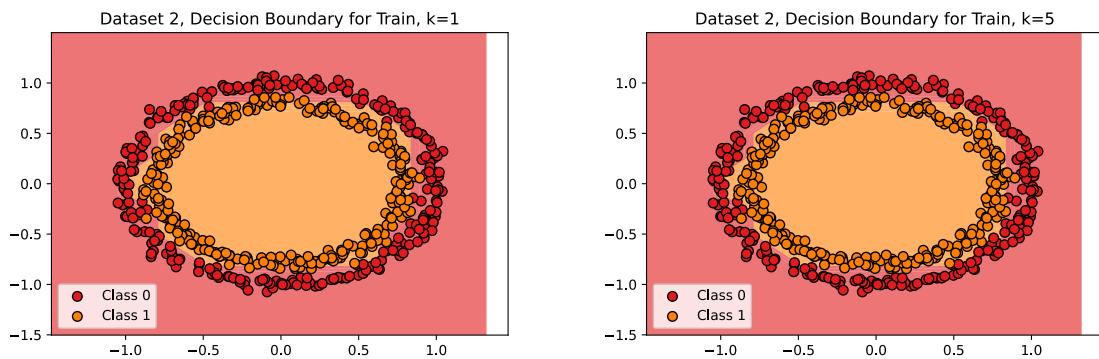
Table 3: KNN Accuracy

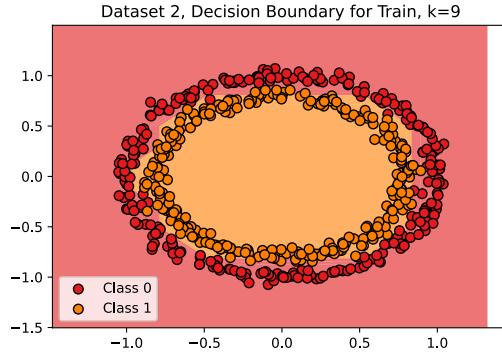
The model using $k = 5$ performs the best on the validation data. All models perform very well for the given classes.

Confusion Matrices



Plots





KNR

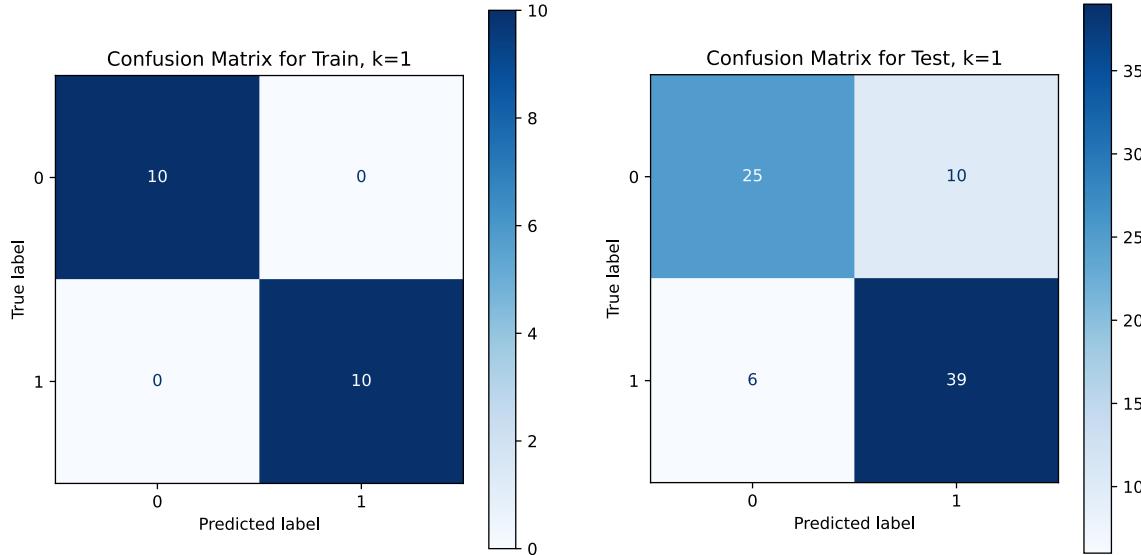
Accuracies

K	Train Accuracy	Validation Accuracy	Testing Accuracy
1	1.000	0.7798	0.7798
3	0.3500	0.6163	0.6375
5	0.8500	0.5400	0.6125

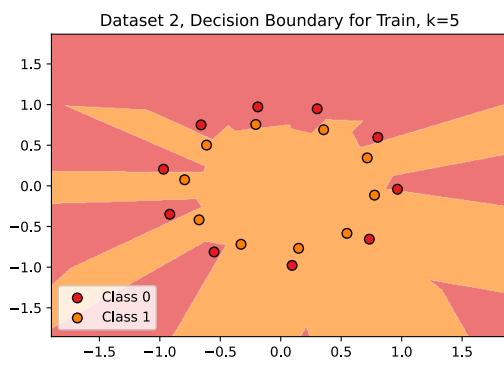
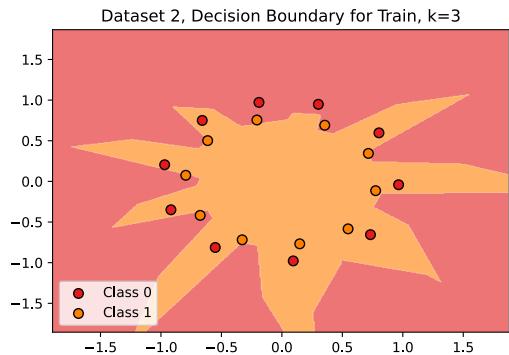
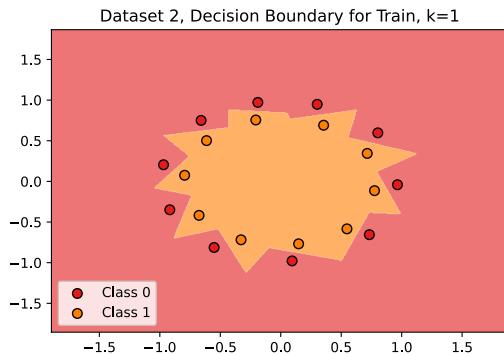
Table 4: KNR Accuracy

The KNR model with $k = 1$ is the most accurate. This is to be expected because under-fitting becomes more likely when clustering.

Confusion Matrices



Plots



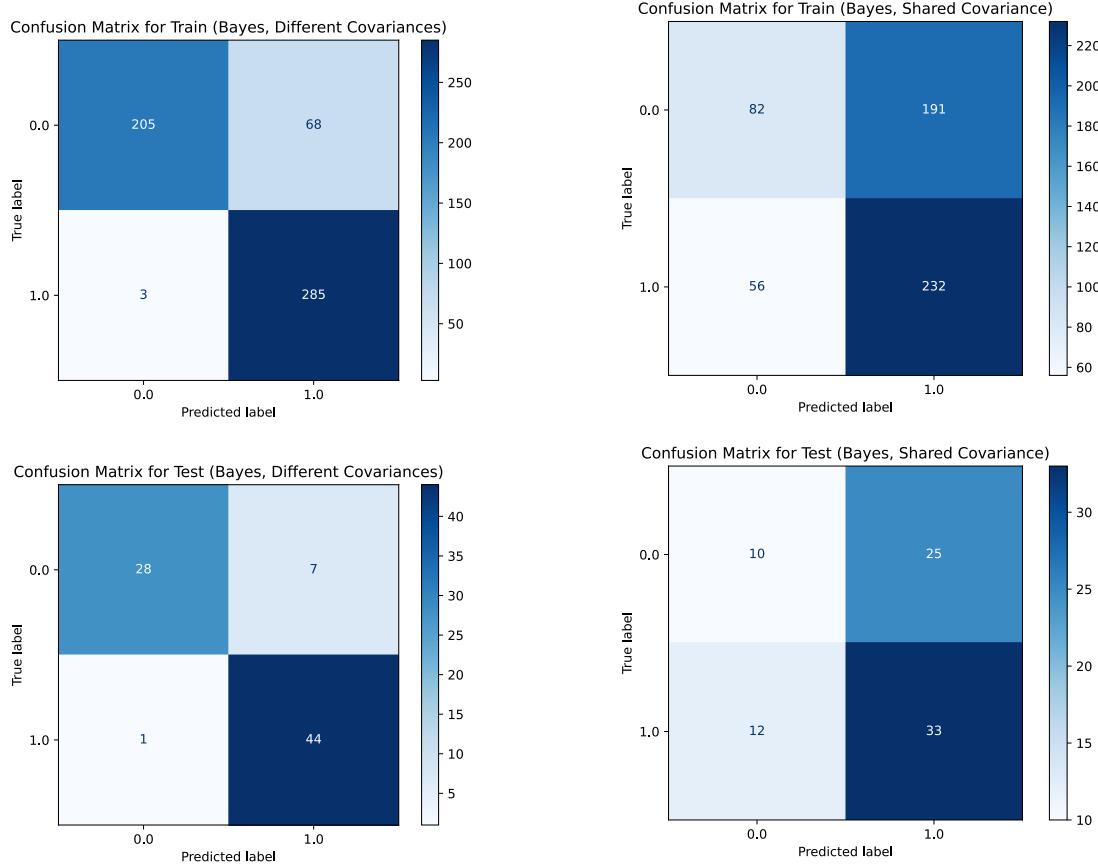
Bayes classifier with Gaussian distribution

Accuracies

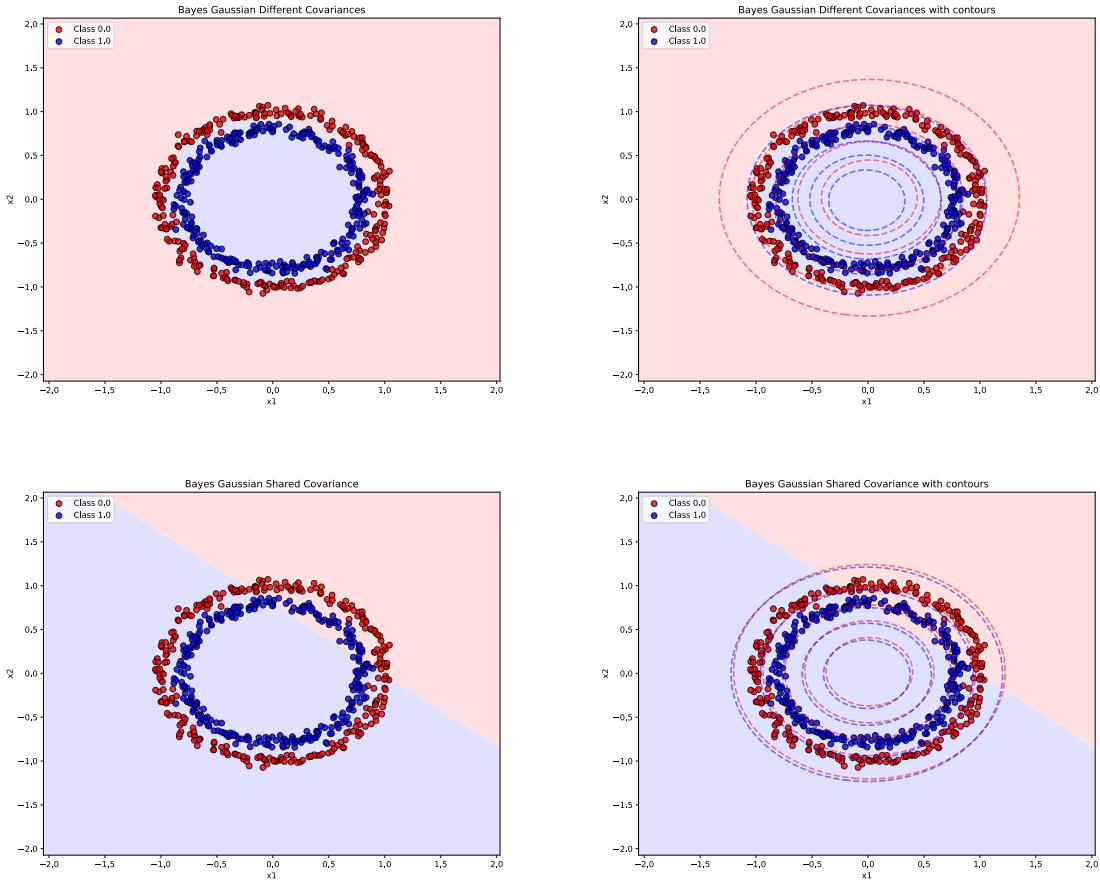
Type	Train Accuracy	Validation Accuracy	Testing Accuracy
With different covariances	0.8734	0.8302	0.9000
With shared covariance	0.5597	0.4591	0.5375

Table 5: Bayes Gaussian Accuracy
Best Model- Bayes Classifier with different covariances

Confusion Matrices



Plots



Naive Bayes classifier with Gaussian distribution

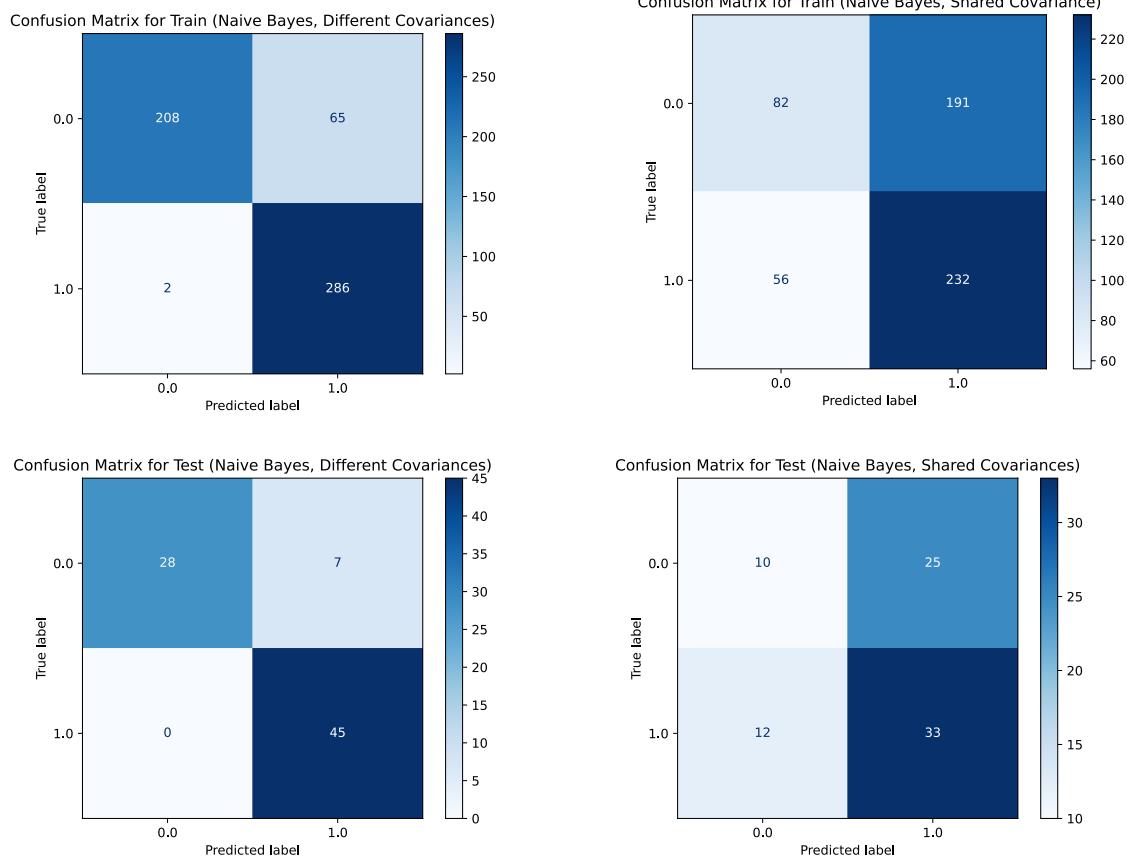
Accuracies

Type	Train Accuracy	Validation Accuracy	Testing Accuracy
With different covariances	0.8806	0.8239	0.9125
With shared covariance	0.5597	0.4591	0.5375

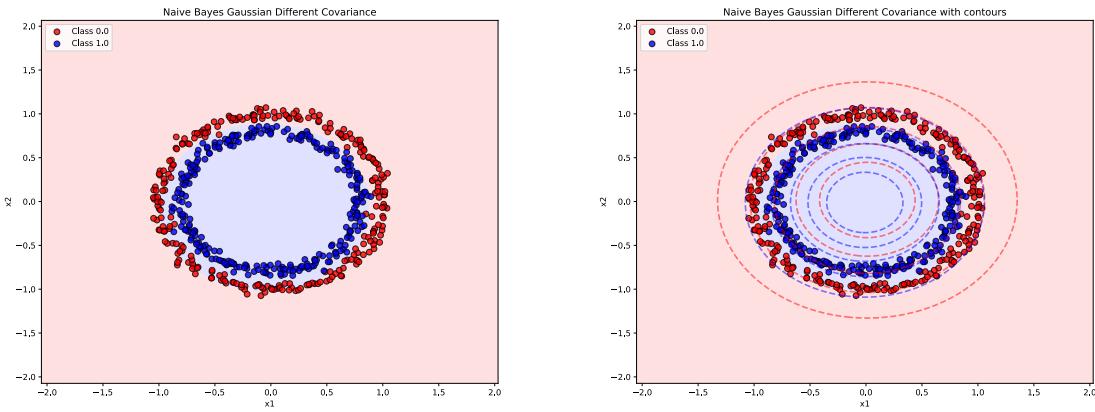
Table 6: Naive Bayes Gaussian Accuracy

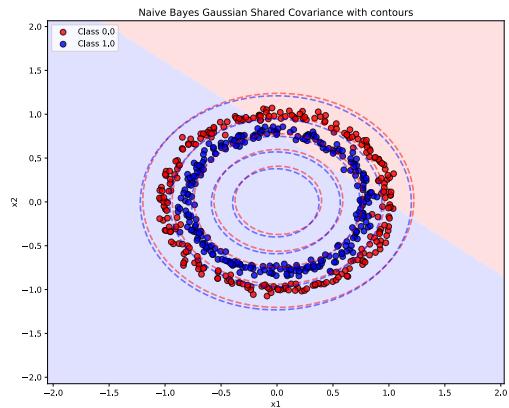
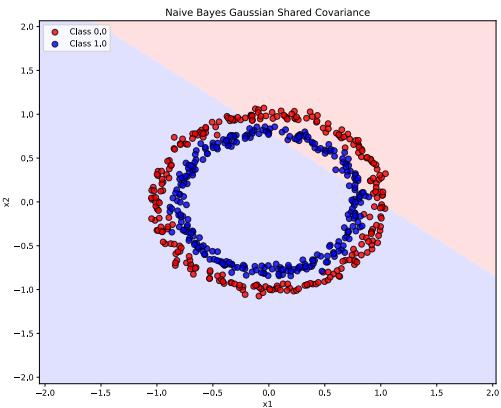
Best Model- Naive Bayes Classifier with different covariances

Confusion Matrices



Plots





Dataset-3

KNN

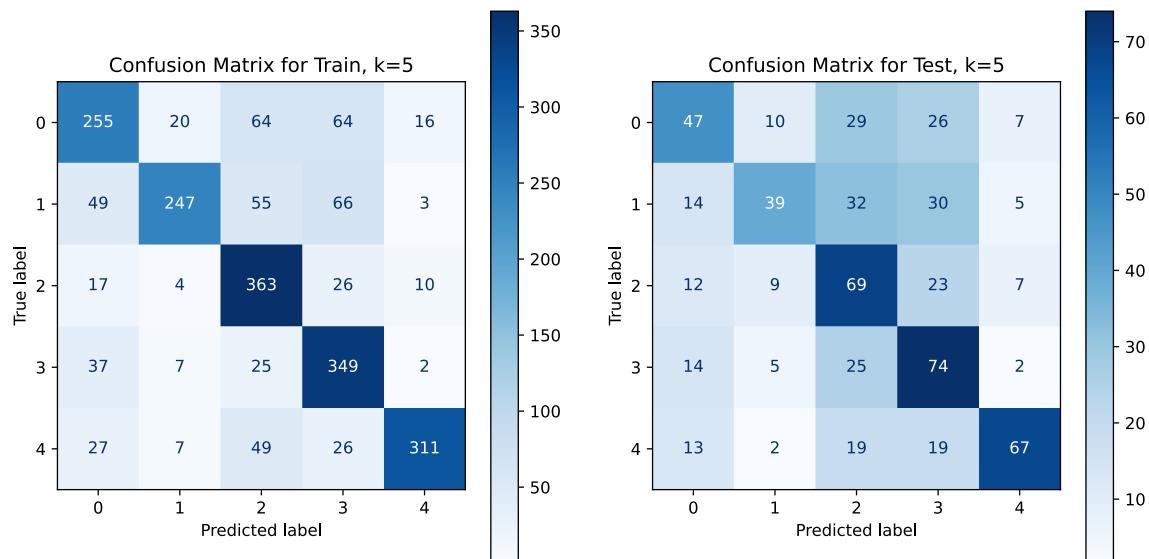
Accuracies

K	Train Accuracy	Validation Accuracy	Testing Accuracy
1	1.000	0.4682	0.4741
5	0.7264	0.511	0.4941
15	0.6112	0.5351	0.5091

Table 7: KNN Accuracy

The KNN model with $k = 5$ performs the best.

Confusion Matrices



KNR

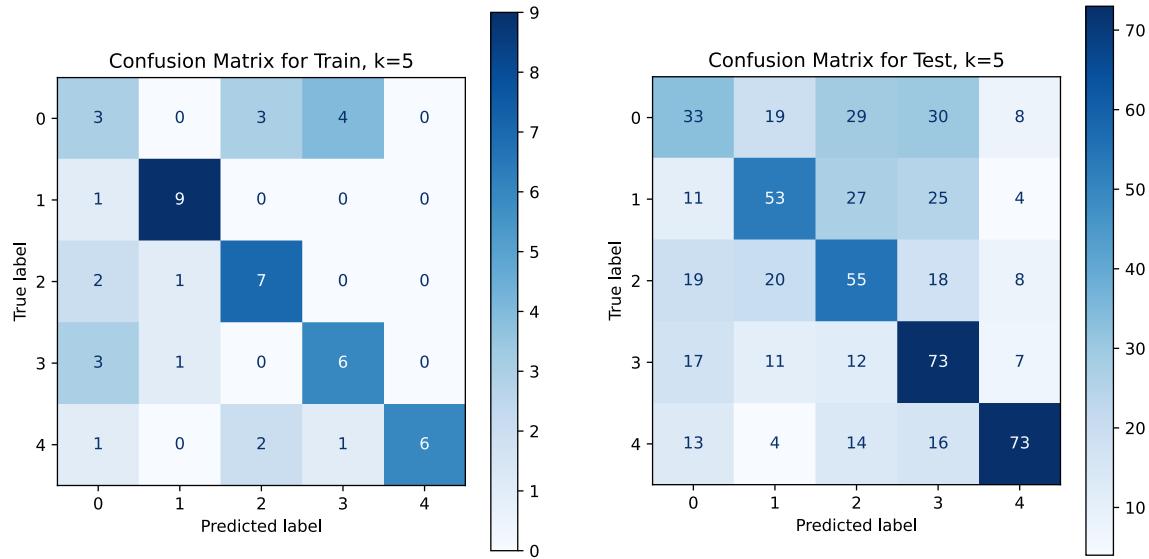
Accuracies

K	Train Accuracy	Validation Accuracy	Testing Accuracy
1	1.000	0.4782	0.4924
5	0.6200	0.5284	0.4791
9	0.6400	0.5083	0.4557

Table 8: KNR Accuracy

The KNR model with $k = 5$ performs the best.

Confusion Matrices



Bayes classifier with Gaussian distribution

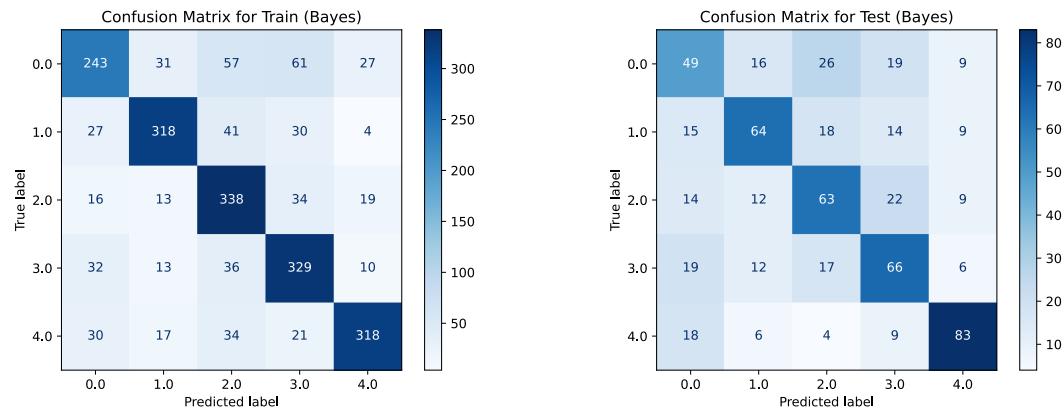
For Dataset 3, Bayes classifier with Gaussian distribution performs better than naive Bayes.

Accuracies

Type	Train Accuracy	Validation Accuracy	Testing Accuracy
With different covariances	0.7365	0.5853	0.5426

Table 9: Bayes Gaussian Accuracy

Confusion Matrices



Naive Bayes classifier with Gaussian distribution

Accuracies

Type	Train Accuracy	Validation Accuracy	Testing Accuracy
With different covariances	0.5665	0.5819	0.5376

Table 10: Naive Bayes Gaussian Accuracy

Confusion Matrices

