For the KNN model on this dataset, all values of K have the same classification accuracies (100%) on train, validation and test datasets, hence all of them are equally good.

The confusion matrices for the best performing model(s) on training and test datasets are given in Tables 2 and 3 respectively.

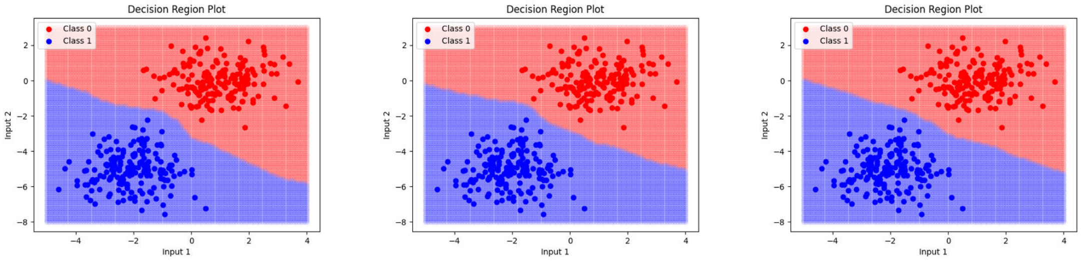
		Actual Class	
		Class 0	**Class 1**
Predicted Class	**Class 0**	182	0
	Class 1	0	178

Table 2: Confusion Matrix of Training Data

		Actual Class	
		Class 0	**Class 1**
Predicted Class	**Class 0**	21	0
	Class 1	0	29

Table 3: Confusion Matrix of Test Data

The decision region plots for all configurations of the model are given in Figure 2.



(a) $K = 1$

(b) $K = 7$

(c) $K = 15$

Figure 2: K Nearest Neighbours Decision Boundaries

1.2.1 Inferences

We infer the following points from the classification accuracies, confusion matrices and decision region plots.

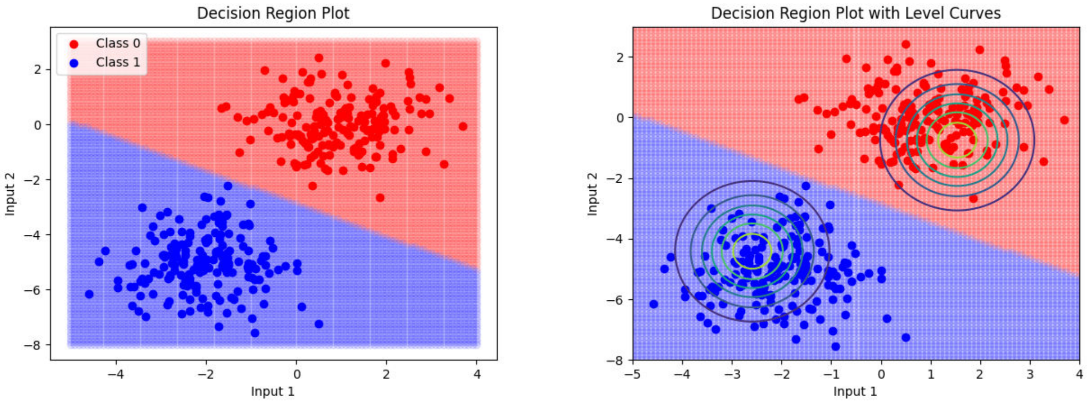
- The KNN method with $K = 1/7/15$ performs exceptionally well on this dataset giving 100% classification accuracy on train, validation and test datasets, Table 1. Therefore, the best performing model gives an accuracy of **100%** on test data.
- The confusion matrices for training and testing data are diagonal as every example in their respective datasets is correctly classified.
- In theory, KNN models do not make explicit specifications about the nature of the decision boundary (can be linear or non-linear). However, as depicted in Figure 2, we can observe an almost linear decision boundary. This occurrence arises from the linear separability inherent in the provided dataset.

1.3 Naive Bayes Classifier

In this section, we employ the Naive Bayes Classifier for Gaussian class conditional density. We consider two cases: one where both classes have the same covariance matrices, and one where they have different covariance matrices.

1.3.1 Model 2(a): Using same covariance matrix for both classes

We take both class covariance matrices to be the same and equal to the average over individual class covariance matrices. The decision region plots with and without level curves are given in Figure 3.



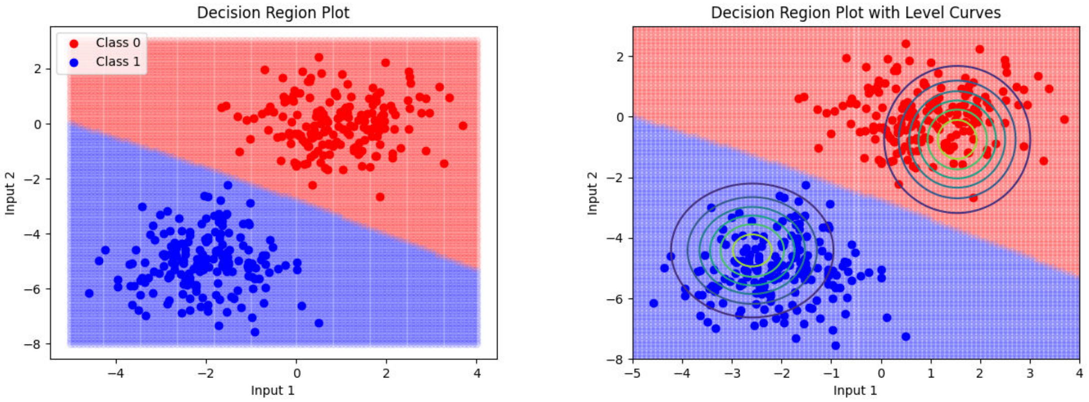
(a) Without Level Curves

(b) With Level Curves

Figure 3: Decision Region for Naive Bayes Classifier with same class covariance matrices

1.3.2 Model 2(b): Using different covariance matrix for each class

In this model, we calculate individual class covariance matrices for each class. The decision region plots with and without level curves are given in Figure 4.



(a) Without Level Curves

(b) With Level Curves

Figure 4: Decision Region for Naive Bayes Classifier with different class covariance matrices

The classification accuracies for both Naive Bayes classifiers are given below in Table 4. The best performing models have been shown in bold.

From the above table, we can see $K = 1, 7, 15$ have the same classification accuracies of 100% on the validation and test datasets. Since $K = 1$ has highest accuracy on the training dataset (100%), we infer that $K = 1$ is the best model.

The confusion matrices for the best performing model on training and test datasets are given in Tables 14 and 15 respectively.

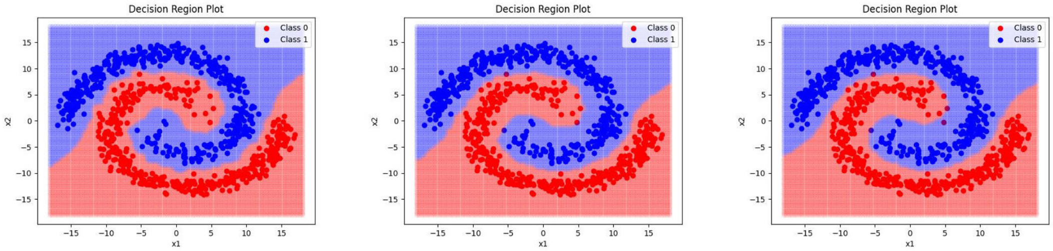
		Actual Class	
		Class 0	**Class 1**
Predicted Class	**Class 0**	411	0
	Class 1	0	431

Table 14: Confusion Matrix of Training Data

		Actual Class	
		Class 0	**Class 1**
Predicted Class	**Class 0**	62	0
	Class 1	0	58

Table 15: Confusion Matrix of Test Data

The decision region plots for various values of K are shown below in Figure 13.



(a) $K = 1$ (b) $K = 7$ (c) $K = 15$

Figure 13: K Nearest Neighbours Decision Boundaries

2.2.1 Inferences

We infer the following points from the classification accuracies, confusion matrices and decision region plots.

- From Table 13, we observe $K = 1, 7, 15$ give perfect scores for validation and testing datasets. However, only the $K = 1$ scenario gives 100% accuracy on the training dataset, therefore this is considered to be the best performing model. The test accuracy on this model is **100%**.

- The confusion matrices, in the case of the best performing model, are diagonal as we get 100% classification accuracies on training, validation and testing datasets.

- As observed from 13, the decision boundaries are **non-linear**. Given that the provided dataset is non-linearly separable, this model performs exceptionally well in accurately classifying the examples.

- We can also observe, from the decision region plots, that increasing K causes more and more red points to seep into the blue decision region, leading to more misclassification.

2.3 Bayes Classifier using Gaussian Distributions

In this section, we employ the Bayes Classifier with single Gaussian class conditional density. We consider full covariance matrices for both classes here. The classification accuracies on training, validation and test data are given in Table 16.

Distribution	**Train Accuracy**	**Validation Accuracy**	**Testing Accuracy**
Single Gaussian	75.53444181	**72.68907563**	70.83333333

Table 16: Bayes Classifier Accuracies

The confusion matrices for this model on training and test datasets are given in Tables 17 and 18 respectively.

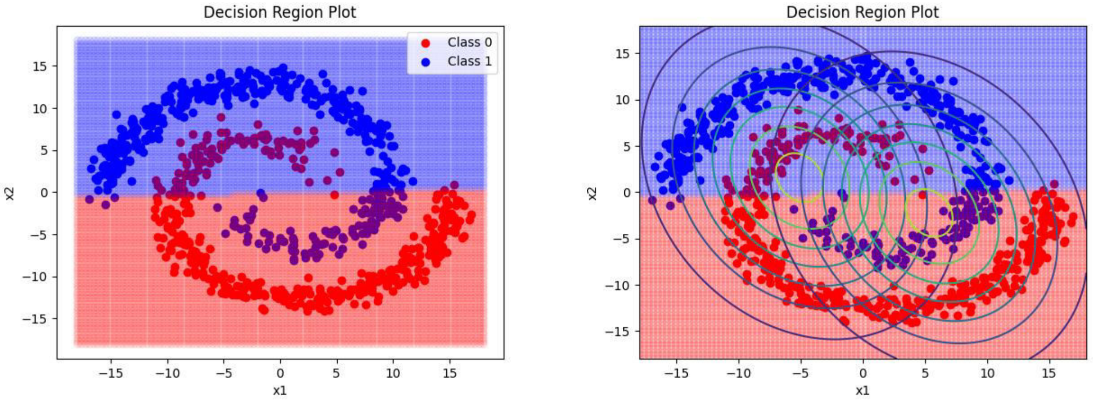
		Actual Class	
		Class 0	**Class 1**
Predicted Class	**Class 0**	312	107
	Class 1	99	324

Table 17: Confusion Matrix of Training Data

		Actual Class	
		Class 0	**Class 1**
Predicted Class	**Class 0**	45	18
	Class 1	17	40

Table 18: Confusion Matrix of Test Data

The decision region plots for this model with and without level curves are given in Figure 14.



(a) Without Level Curves

(b) With Level Curves

Figure 14: Gaussian Bayes Classifier Decision Boundaries

3 Dataset 2

We are given an image dataset for 5 classes. The dimension of each feature vector is 81. There are 3 different datasets of the same nature: training, validation and test datasets.

- **Training Dataset:** 3500 examples
- **Validation Dataset:** 1000 examples
- **Testing Dataset:** 500 examples

Each data point can be classified into one of the 5 classes.

3.1 K Nearest Neighbours

We employ the K Nearest Neighbours algorithm to build a classification model for the given dataset. We test the performance of the model on the dataset for K = 1, 7, 15 and document the classification accuracies in Table 31. The best performing model on validation data is shown in bold.

K	**Train Accuracy**	**Validation Accuracy**	**Test Accuracy**
1	100	46	43
7	60.08571429	48.8	47.4
15	56.28571429	**50.3**	46.4

Table 31: KNN Accuracies

From the above table, we can see $K = 15$ has the best classification accuracy of 50.3% on the validation dataset. Hence, we infer that $K = 15$ is the best model.

The confusion matrix on the training dataset for the best performing model is given below.

		Actual Class				
		Class 1	**Class 2**	**Class 3**	**Class 4**	**Class 5**
Predicted Class	**Class 1**	304	91	32	53	29
	Class 2	47	229	23	67	15
	Class 3	240	262	588	260	96
	Class 4	62	98	26	306	17
	Class 5	47	20	31	14	543

The confusion matrix on the test dataset for the best performing model is given below.

		Actual Class				
		Class 1	**Class 2**	**Class 3**	**Class 4**	**Class 5**
Predicted Class	**Class 1**	31	9	7	7	3
	Class 2	7	16	3	14	4
	Class 3	45	48	74	41	16
	Class 4	12	20	10	35	1
	Class 5	5	7	6	3	76