REPORT

ASSIGNMENT – 1

PATTERN RECOGNITION (CS669)

SUBMITTED BY:

Abhishek Tandon

Roll Number: S22002

MTech (by Research) - Intelligent Systems

School of Computing and Electrical Engineering

LINKS

- Google Colab - Question 1
- Google Colab - Question 2
- View on GitHub

TABLE OF CONTENTS

QUESTION 1

1.1. PROBLEM STATEMENT

We are required to perform Speech Activity Detection (SAD) on a given sequence of signal frames and classify each frame as speech or non-speech signal. There are two types of 1-D features provided: short-time energy (STE), and Mel-filterbank energy (MEL). We have to determine which of these features are better at correctly detecting speech.

1.2. EXPECTED RESULTS

- Estimated distribution of the features using unimodal Gaussian with sample mean and sample variance as parameters.
- Plot ROC Curve for each of the 1D features provided: STE and MEL energy.
- Compare ROC Curves for STE and MEL energy and state which feature is better to classify a signal frame.

1.3.SOLVING THE PROBLEM

1.3.1. DESCRIPTION OF DATA

The given dataset has 256 samples each for ST-Energy and MEL Energy for two classes: speech and non-speech

Feature – ST-Energy	Number of samples
Class 1: Speech	111
Class 2: Non-Speech	145
Total	256

Table 1.1 Training data - ST-Energy

Feature – MEL Energy	Number of samples
Class 1: Speech	111
Class 2: Non-Speech	145
Total	256

Table 1.2 Training data - MEL Energy

1.3.2. SAMPLE MEAN AND SAMPLE VARIANCE

Parameters in Gaussian Distribution: sample mean and sample variance are required to estimate the normal curve for training data so as to later compute P(x given class) or the likelihood probability.

Formulae used are given as follows:

- $$mean: \ \bar{x} = \frac{\sum_1^n x_i}{N}$$
- $$variance: \ \sigma^2 = \frac{\sum_1^n (x_i - \bar{x})^2}{N}$$

ST-Energy (Train)	Sample Mean	Sample Variance
Class 1: Speech	0.12043	0.02768
Class 2: Non-Speech	0.05309	0.00129

Table 1.1 Mean and Variance for STE

MEL-Energy (Train)	Sample Mean	Sample Variance
Class 1: Speech	0.61005	0.02287
Class 2: Non-Speech	0.45865	0.00518

Table 1.1 Mean and Variance for MEL

1.3.3. ESTIMATED GAUSSIAN DISTRIBUTION

Likelihood is assumed to be taking values from unimodal gaussian distribution with mean and variance as the parameters.

$$P(x|class) \sim N(\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left[\frac{-1}{2}\left(\frac{x-\bar{x}}{\sigma}\right)^2\right]$$

1.3.4. FIGURES – GAUSSIAN DISTRIBUTION

We have plotted the estimated curve each for STE energy and MEL energy using their respective sample mean and variances for each class. We get two curves in the same plot each depicting distribution of speech and non-speech training data.

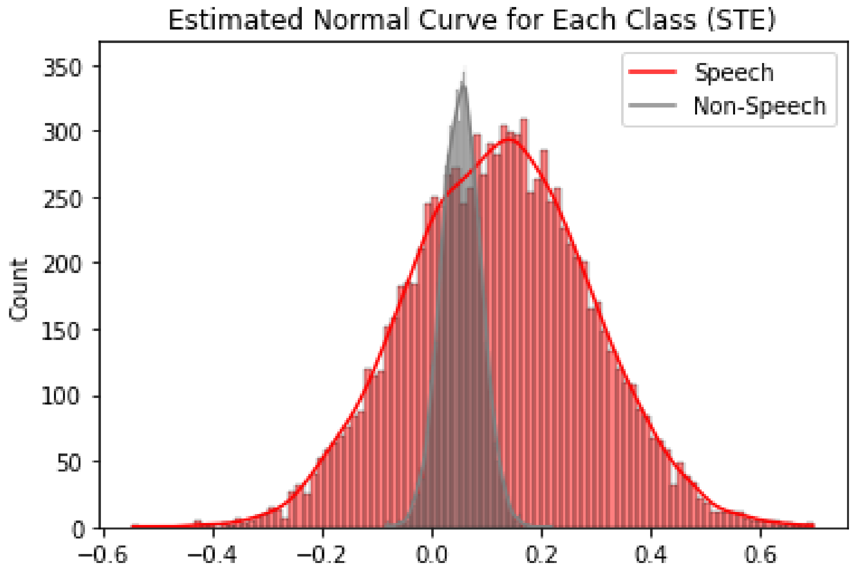

Fig. 1.1. Estimated Normal Curve for STE

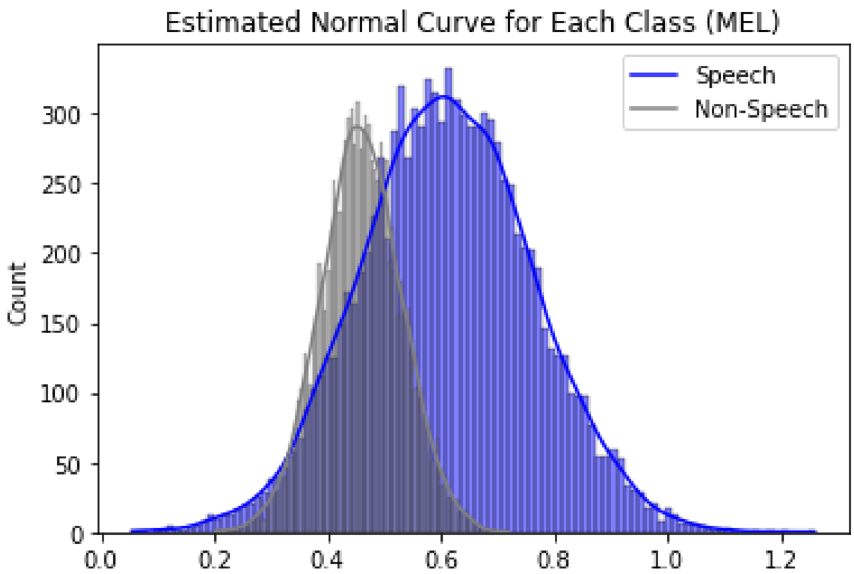

Fig. 1.2. Estimated Normal Curve for MEL

1.3.5. BAYES' THEOREM

The Bayes' formula applied in the context of speech activity detection is used so as to determine with what probability a given frame energy is a speech signal.

$$P(speech|x) = \frac{P(x|speech)\,P(speech)}{P(x|speech)\,P(speech) + P(x|nonSpeech)\,P(nonSpeech)};\; where\;\; x\; = test\,data$$

Here,

- $P(speech|x)$ is the *posterior probability* to determine how likely is the given frame energy a speech signal
- $P(x|speech)$ is the *likelihood* that assumes values from a gaussian distribution with $\bar{x}$ and σ^2 as parameters.
- $P(speech)$ and $P(nonSpeech)$ is the *prior probability* of a given class that is calculated as follows:

$$P(class) = \frac{Number\ of\ samples\ in\ class\ i}{Total\ number\ of\ samples}$$

- $P(x|speech)\,P(speech) +\ P(x|nonSpeech)\,P(nonSpeech)$ is the evidence, which acts as a normalising term in the equation

1.3.6. TESTING THE MODEL

Our test data consists of 275 frame energy values to test our model based on ST-Energy and MEL energy features separately.

1.3.7. PLOTTING THE ROC CURVE

In order to plot the ROC curve for the samples, we need to determine the *True Positive Rate* (TPR) as well as *False Positive Rate* (FPR), which are given as follows:

$$TPR = \frac{True\ Positives}{True\ Positives + False\ Negatives}$$

$$FPR = \frac{False\ Positives}{False\ Positives + True\ Negatives}$$

1.4. RESULTS

Figures for each ROC Curve are as depicted below. They are plotted with the help of the True Positive Rate and False Positive Rate as mentioned in 1.3.7.

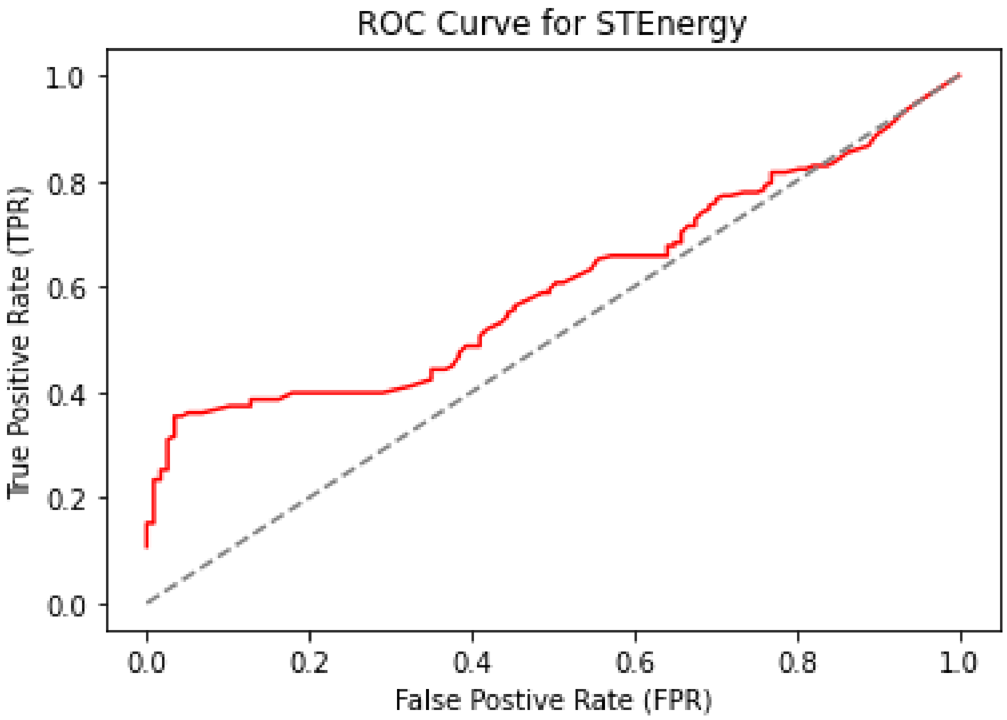

Fig. 1.3. ROC Curve for ST-Energy

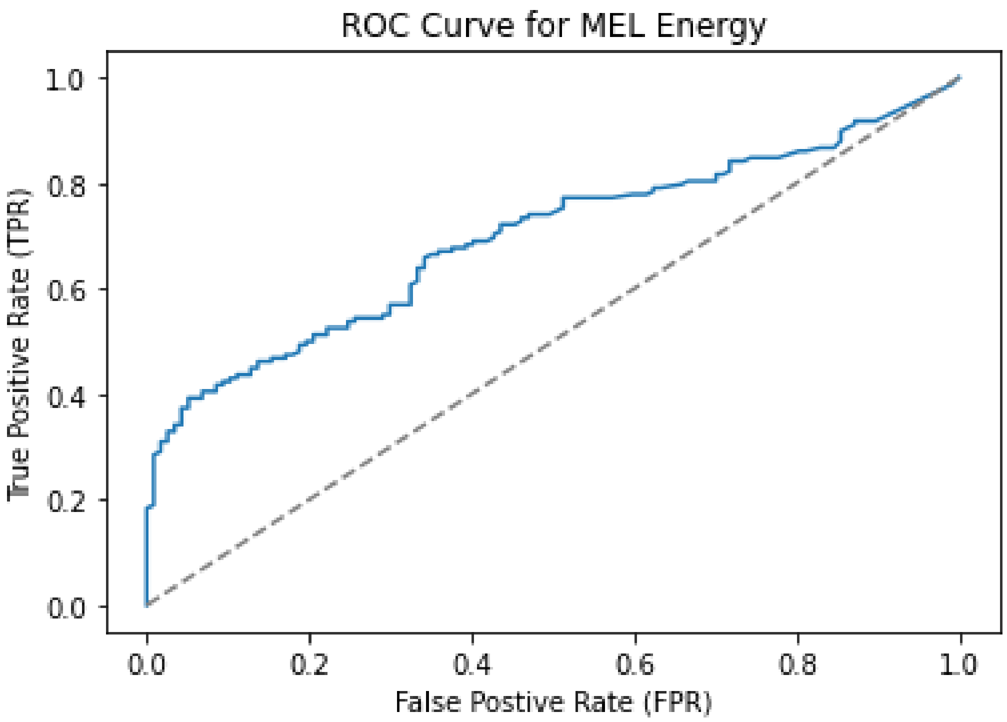

Fig. 1.4. ROC Curve for ST-Energy

1.5. COMPARISON BETWEEN ROC CURVES (MEL V/S ST-ENERGY)

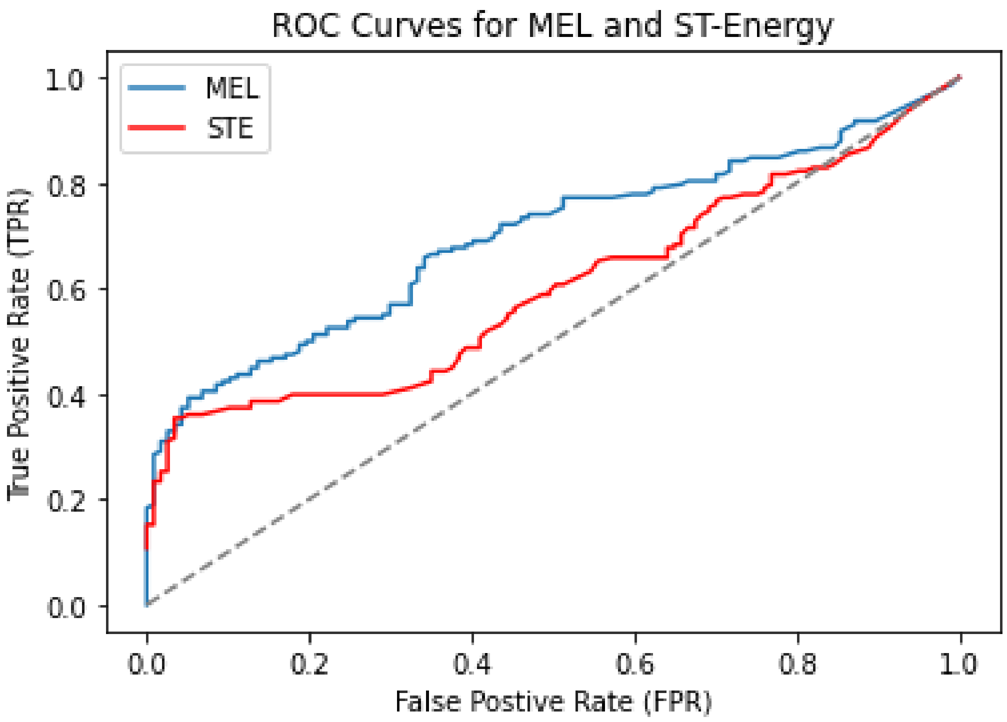

Fig. 1.5. ROC Curves for MEL and STE energy features in the same plot

1.6. CONCLUSION

MEL Energy gives better results as compared to ST-Energy since the ROC curve clearly indicates any given pair of (FPR, TPR) has a greater value for the MEL curve as compared to the ST-Energy curve. Therefore, MEL energy is better feature to develop a speech activity detection system.

QUESTION 2

2.1. PROBLEM STATEMENT

We are required to design a Bayes classifier with Gaussian distribution of data for each of the 3 classes (in two separate datasets). The first data set has linearly separable data and the second classifier has non-linearly separable data. We need to repeat the process to develop 4 classifiers with by varying the covariance matrix used to build the classifier.

2.2. DESCRIPTION OF DATA

Total Data – shape of train data and class wise shape

Linearly Separable Data	
	Number of samples
Class 1	500
Class 2	500
Class 3	500
Total	1500

Table 2.1 Training data – Linearly Separable

Non-Linearly Separable Data	
	Number of samples
Class 1	500
Class 2	500
Class 3	500
Total	1500

Table 2.2 Training data – Non-linearly Separable

We have further split this data randomly into train and test datasets (50% each). Each dataset has two features; hence this is multivariate classification problem.

The following scattered plots depict the nature of the two datasets – linearly separable and non-linearly separable.

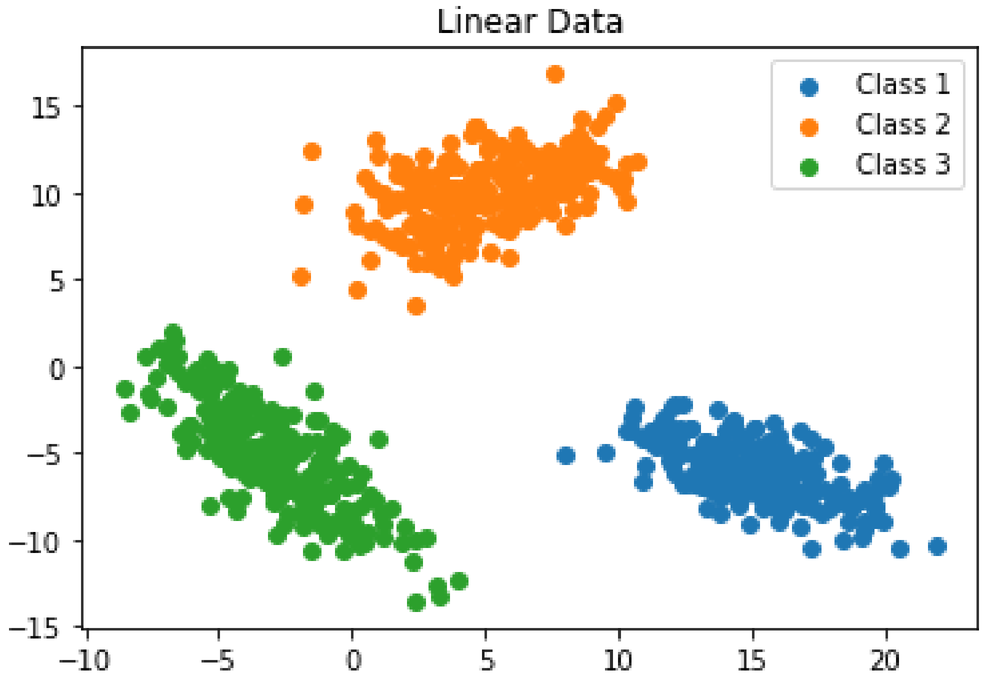

Fig 2.1 Training data – Linearly Separable

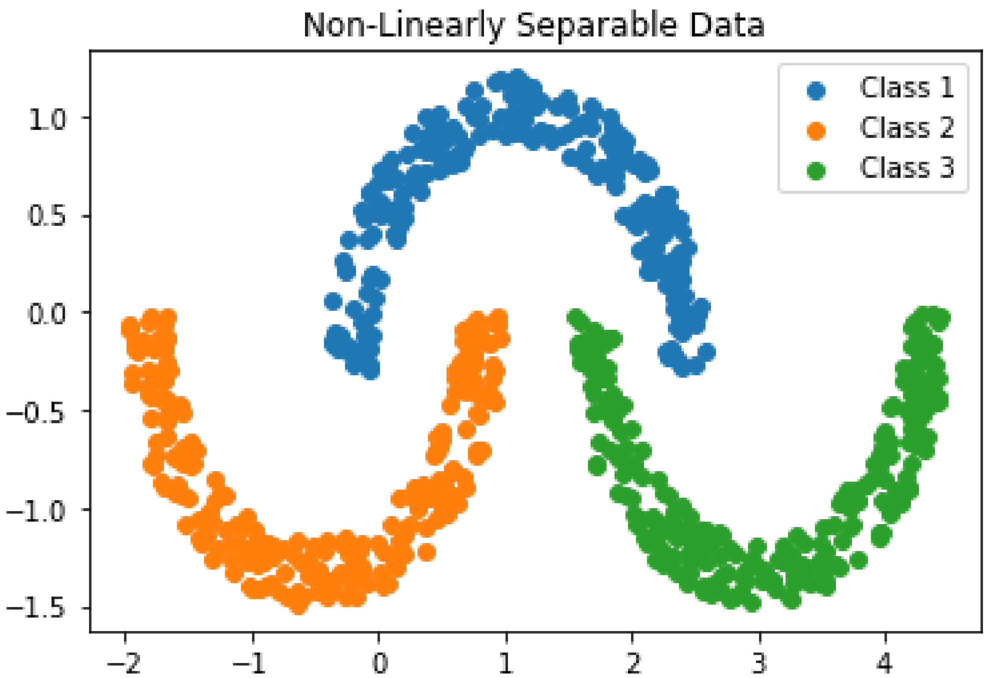

Fig 2.2 Training data – Non-linearly Separable

2.3. SAMPLE MEAN AND SAMPLE VARIANCE

Sample mean and variance have been calculated using the following formulae:

- $$mean: \ \bar{x} = \frac{\sum_1^n x_i}{N}$$
- $$variance: \ \sigma^2 = \frac{\sum_1^n (x_i - \bar{x})^2}{N}$$

2.4. MULTIVARIATE GAUSSIAN DISTRIBUTION

Since features are two dimensional, the gaussian distribution with parameters μ (mean) and Σ (covariance matrix) is given as:

$$P(x|class) \sim N(\mu, \Sigma) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp\left[\frac{-1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right]$$

2.5. BAYES' THEOREM

We have used Bayes' Theorem to develop the classifier. It is described as follows:

$$P(c1|x) = \frac{P(x|c1)P(c1)}{P(x|c1)P(c1) + P(x|c2)P(c2) + P(x|c3)P(c3)}$$

$$P(c2|x) = \frac{P(x|c2)P(c2)}{P(x|c1)P(c1) + P(x|c2)P(c2) + P(x|c3)P(c3)}$$

$$P(c3|x) = \frac{P(x|c3)P(c3)}{P(x|c1)P(c1) + P(x|c2)P(c2) + P(x|c3)P(c3)}$$

- $P(ci|x)$ is the *posterior probability* to determine how likely x belongs to a particular class i
- $P(x|ci)$ is the *likelihood* that assumes values from a gaussian distribution with μ and Σ as parameters.
- $P(ci)$ is the *prior probability* of a given class that is calculated as follows:

$$P(class) = \frac{Number\ of\ samples\ in\ class\ i}{Total\ number\ of\ samples}$$

- $P(x|c1)\,P(c1) + P(x|c2)\,P(c2) + P(x|c3)\,P(c3)$ is the evidence = $P(x)$
- The highest probability amongst $P(c1|x)$, $P(c2|x)$, and $P(c3|x)$ is where the data point x is considered to be belong to the respective class

2.6. EVALUATION METRICS:

The performance metrics used to evaluate our classifiers are as follows:

$$Accuracy = \frac{True\ Positive + True\ Negative}{True\ Positives +\ True\ Negative + False\ Positive + False\ Negative}$$

$$Precision = \frac{True\ Positive}{(True\ Positive + False\ Positive)}$$

$$Recall = \frac{True\ Positive}{(True\ Positive + False\ Negative)}$$

$$F1\ Score = 2 * \frac{(Precision\ *\ Recall)}{(Precision\ +\ Recall)}$$

2.7. COVARIANCE MATRIX

The covariance matrix used here for 2 features, can be given as

$$\Sigma = \begin{pmatrix} var(x_1) & cov(x_1, x_2) \\ cov(x_1, x_2) & var(x_2) \end{pmatrix}$$

where x_1 and x_2 are the features or the dimensions of the dataset and the covariance between the two features is given by

$$cov(x_1, x_2) = \frac{\sum(x - \bar{x})(y - \bar{y})}{N}$$

2.8. DESCRIPTION OF THE CLASSIFIERS (C1, C2, C3, C4)

The 4 types of classifiers to be developed are:

- **C1**: Covariance for all classes is $I\sigma^2$. Average of the sample variances for all dimensions and for all classes from the training data is being called as σ^2 .

$$\Sigma = I\sigma^2,\ where$$

$$\sigma^2 = \frac{1}{3}(\sigma_1^2 + \sigma_2^2 + \sigma_3^2);\ \ \sigma_i = variance\ of\ class_i\ for\ all\ dimensions$$

- **C2:** Full but equal covariance for all classes, Σ. Average of the sample covariance matrix from all classes in the train data is used as Σ.

$$\Sigma = \frac{1}{3}(\Sigma_1 + \Sigma_2 + \Sigma_3)$$
$$where\ \Sigma_i\ is\ the\ sample\ covariance\ matrix\ from\ a\ class$$

- **C3:** Diagonal covariance matrix, distinct for each class is used. Variances from the sample covariance matrix for each class is used as $var(x_i)$

$$\Sigma_i = \begin{pmatrix} var(x_1) & 0 \\ 0 & var(x_2) \end{pmatrix}$$

- **C4:** Sample covariance matrix (full) is used, distinct for each class.

$$\Sigma_i = \begin{pmatrix} var(x_1) & cov(x_1, x_2) \\ cov(x_1, x_2) & var(x_2) \end{pmatrix}$$

2.9. RESULTS

The plots depicting the decision boundaries drawn of the 4 different Bayesian classifiers as well their performance (Accuracy, Precision, Recall and F1-Score) designed for linearly separable and non-separable data can be seen in 2.9.1. and 2.9.2 respectively.

2.9.1. LINEARLY SEPARABLE DATA

- **CLASSIFIER 1**

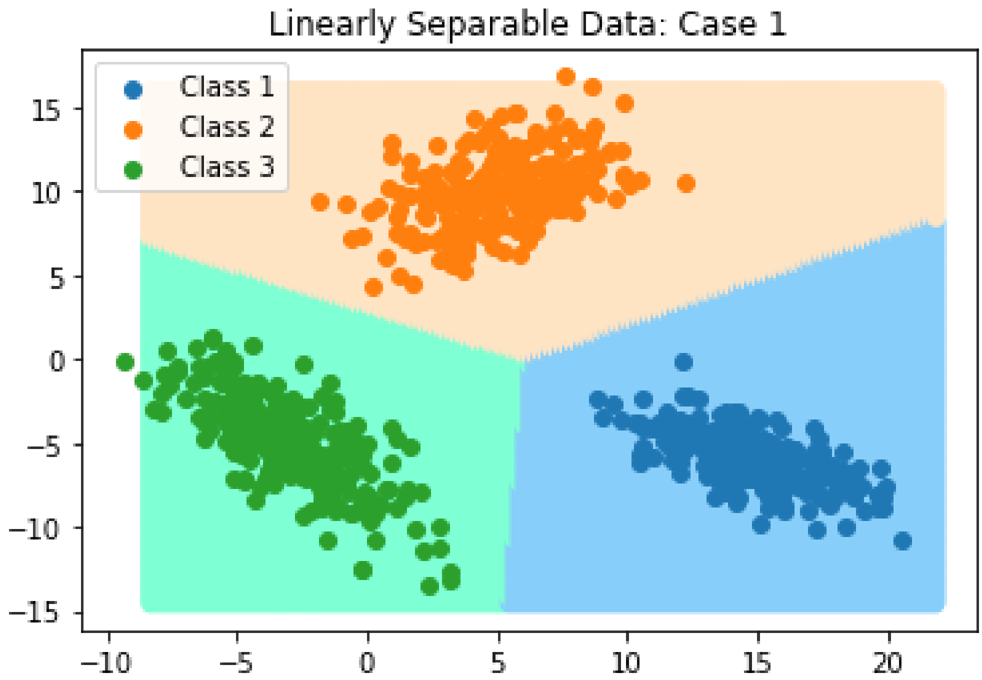

Fig. 2.3. Decision Boundaries – Classifier 1 (Linearly Separable Data)

Metric	Value
Accuracy	1.0
Precision	1.0
Recall	1.0
F1 Score	1.0

Table 2.3. Evaluation Metrics – Classifier 1 (Linearly Separable Data)

- **CLASSIFIER 2**

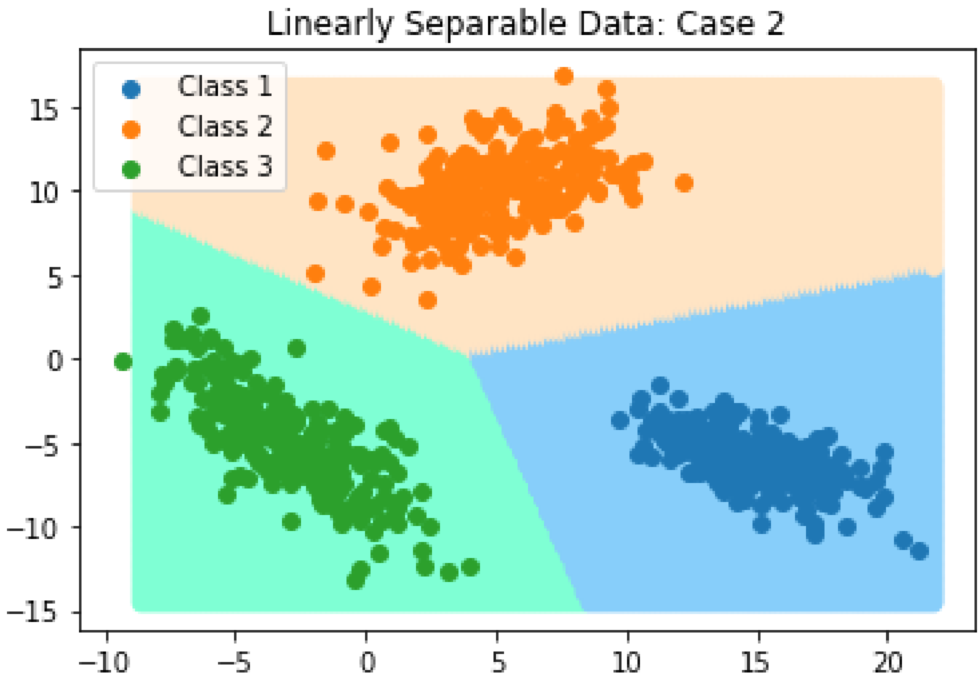

Fig. 2.4. Decision Boundaries – Classifier 2 (Linearly Separable Data)

Metric	Value
Accuracy	1.0
Precision	1.0
Recall	1.0
F1 Score	1.0

Table 2.4. Evaluation Metrics – Classifier 2 (Linearly Separable Data)

- **CLASSIFIER 3**

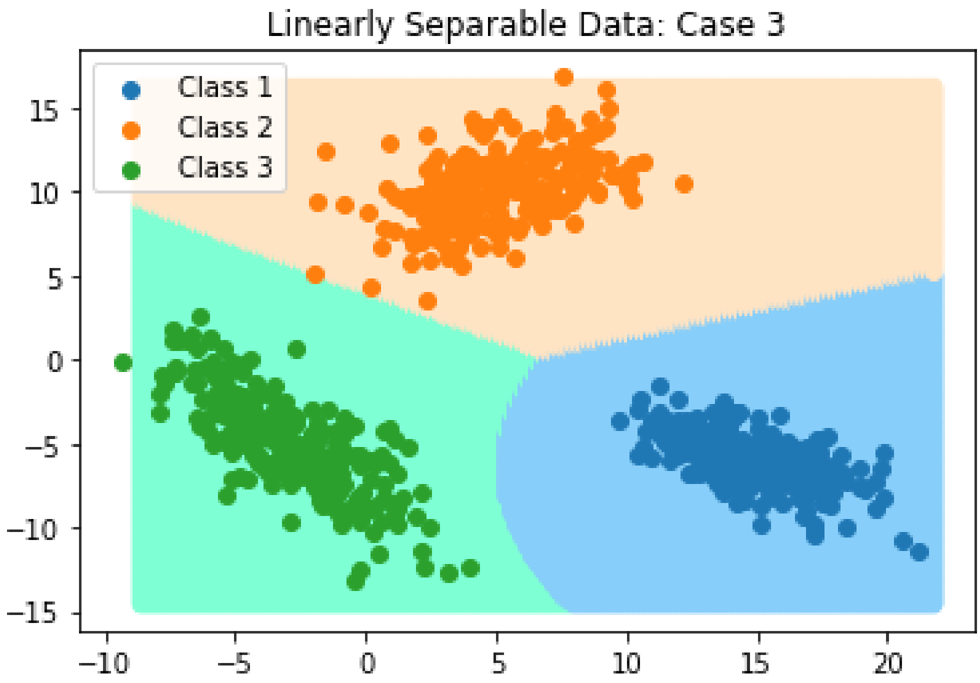

Fig. 2.5. Decision Boundaries – Classifier 3 (Linearly Separable Data)

Metric	Value
Accuracy	1.0
Precision	1.0
Recall	1.0
F1 Score	1.0

Table 2.5. Evaluation Metrics – Classifier 3 (Linearly Separable Data)

- **CLASSIFIER 4**

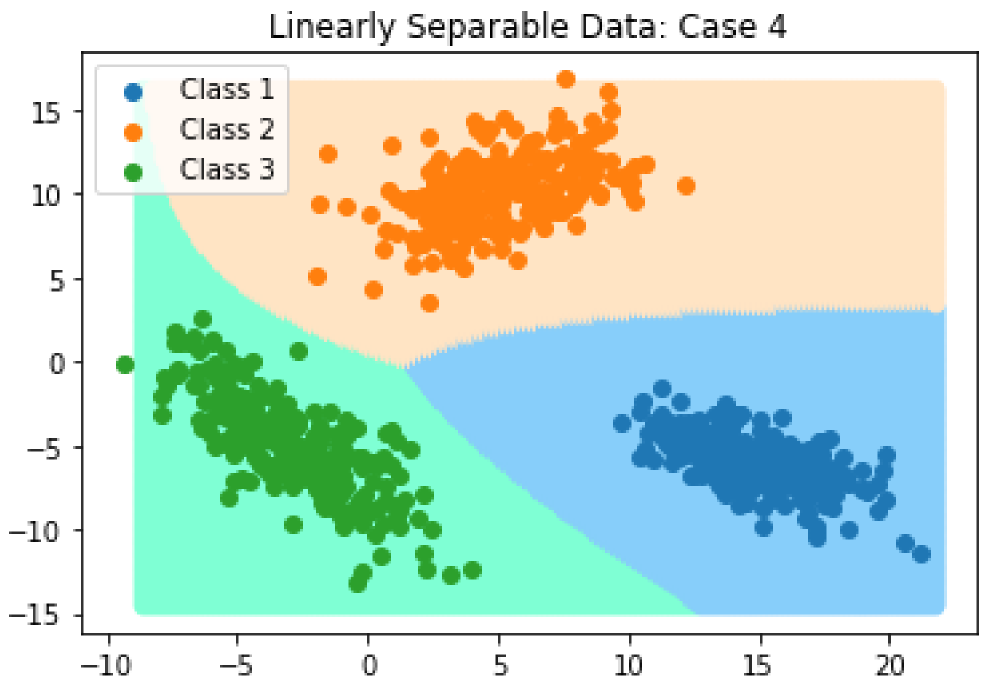

Fig. 2.6. Decision Boundaries – Classifier 4 (Linearly Separable Data)

Metric	Value
Accuracy	1.0
Precision	1.0
Recall	1.0
F1 Score	1.0

Table 2.6. Evaluation Metrics – Classifier 4 (Linearly Separable Data)

2.9.2. NON-LINEARLY SEPARABLE DATA

- **CLASSIFIER 1**

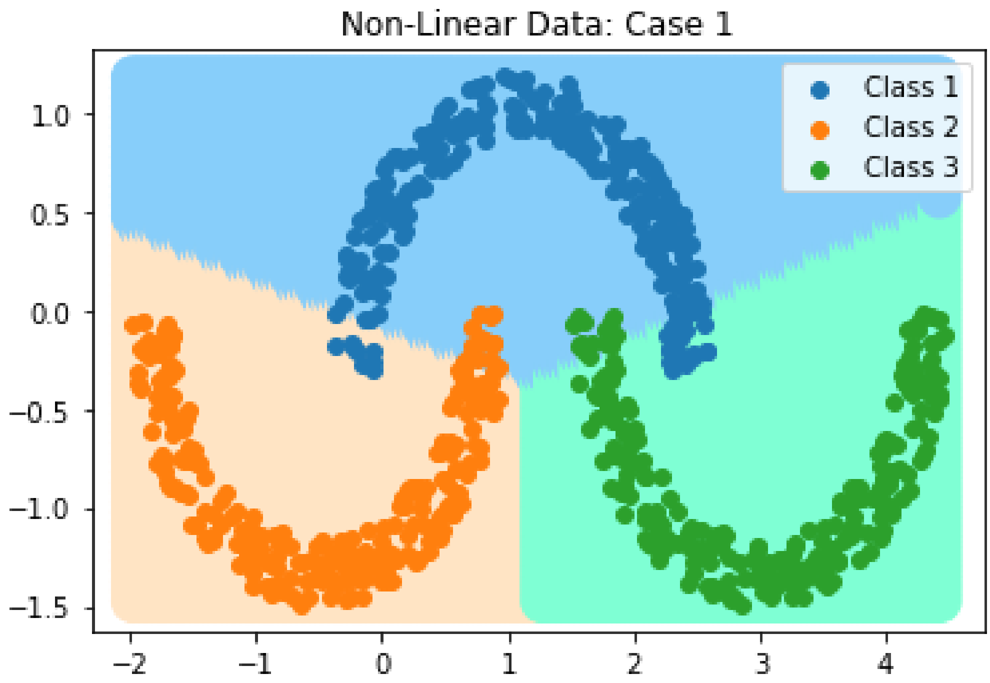

Fig. 2.7. Decision Boundaries – Classifier 1 (Non-Linearly Separable Data)

Metric	Value
Accuracy	0.94044
Precision	0.91116
Recall	0.91066
F1 Score	0.91086

Table 2.7. Evaluation Metrics – Classifier 1 (Non-Linearly Separable Data)

- **CLASSIFIER 2**

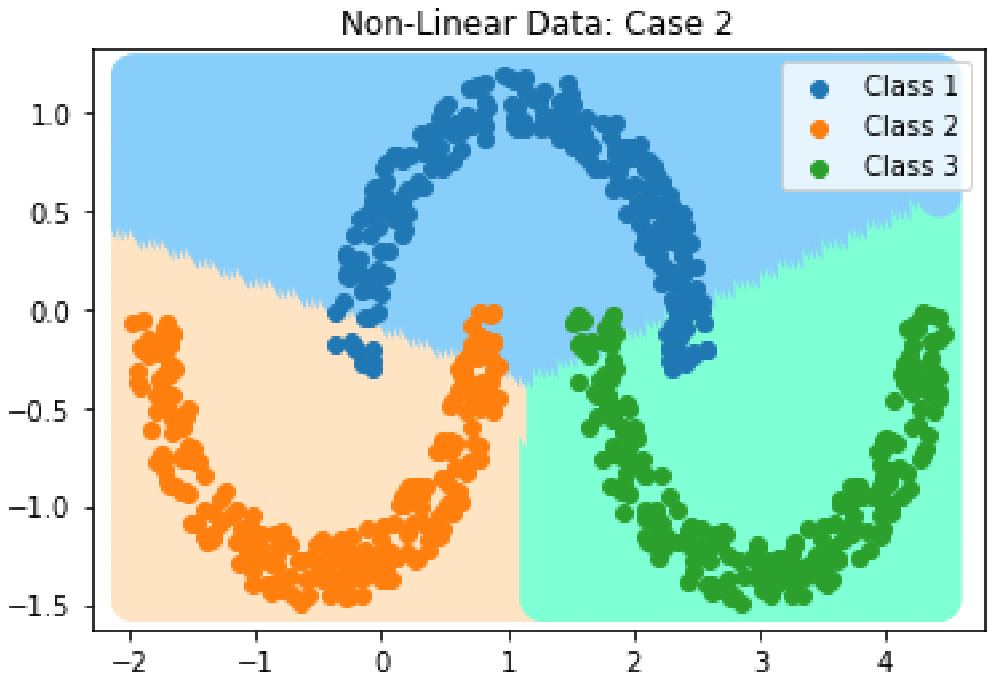

Fig. 2.8. Decision Boundaries – Classifier 2 (Non-Linearly Separable Data)

Metric	Value
Accuracy	0.93955
Precision	0.90972
Recall	0.90933
F1 Score	0.90948

Table 2.8. Evaluation Metrics – Classifier 2 (Non-Linearly Separable Data)

- **CLASSIFIER 3**

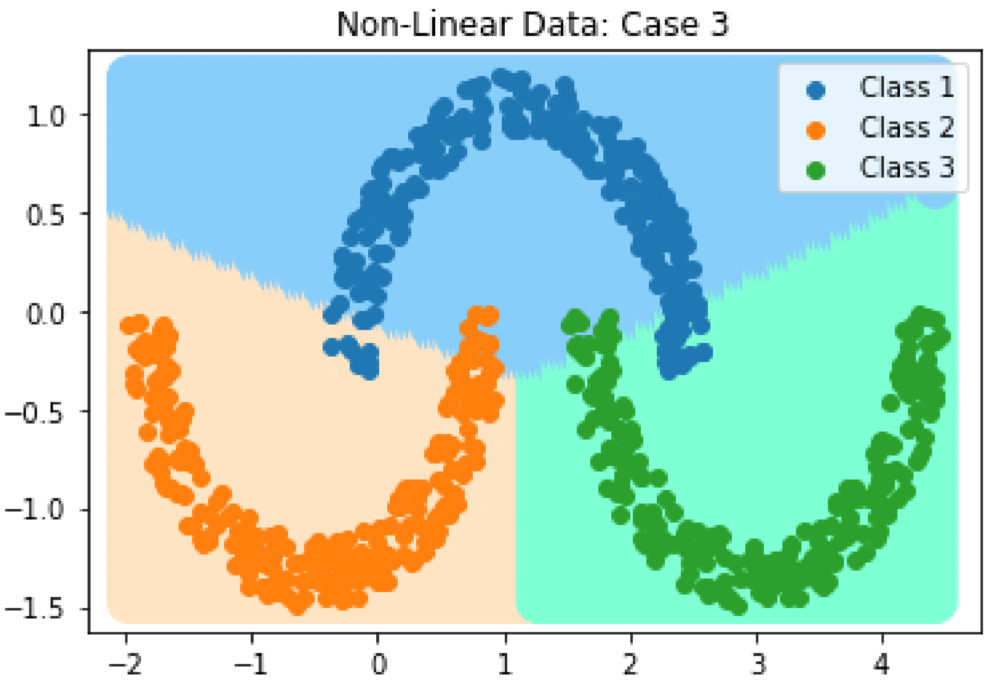

Fig. 2.9. Decision Boundaries – Classifier 3 (Non-Linearly Separable Data)

Metric	Value
Accuracy	0.94222
Precision	0.91361
Recall	0.91333
F1 Score	0.91345

Table 2.9. Evaluation Metrics – Classifier 3 (Non-Linearly Separable Data)

- **CLASSIFIER 4**

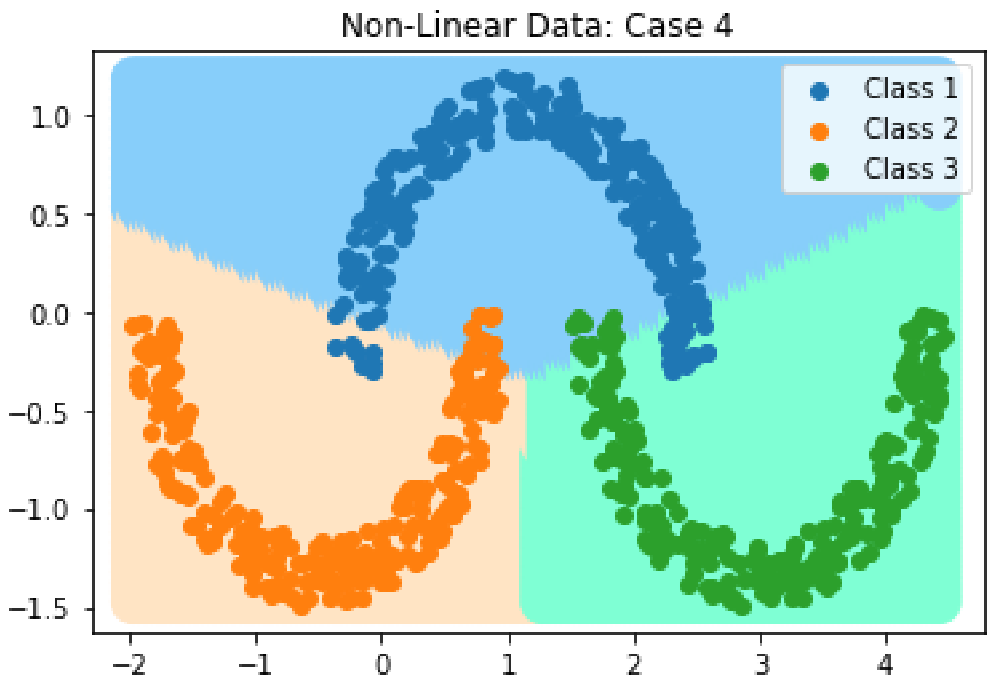

Fig. 2.10. Decision Boundaries – Classifier 4 (Non-Linearly Separable Data)

Metric	Value
Accuracy	0.93866
Precision	0.90819
Recall	0.90799
F1 Score	0.90801

Table 2.10. Evaluation Metrics – Classifier 4 (Non-Linearly Separable Data)

2.10. PERFORMANCE EVALUATION

We can summarize our classifiers each for linearly separable and non-linearly separable data in Table 2.11 and Table 2.12 respectively.

- **LINEARLY SEPARABLE DATA**

Classifier	Accuracy	Precision	Recall	F1 Score
C1	1.0	1.0	1.0	1.0
C2	1.0	1.0	1.0	1.0
C3	1.0	1.0	1.0	1.0
C4	1.0	1.0	1.0	1.0

Table 2.11. Evaluation Metrics – All Classifiers (Linearly Separable Data)

- **NON-LINEARLY SEPARABLE DATA**

Classifier	Accuracy	Precision	Recall	F1 Score
C1	0.94044	0.91116	0.91066	0.91086
C2	0.93955	0.90972	0.90933	0.90948
C3	0.94222	0.91361	0.91333	0.91345
C4	0.93866	0.90819	0.90799	0.90801

Table 2.12. Evaluation Metrics – All Classifiers (Non-Linearly Separable Data)

2.11. CONCLUSION

We developed 4 classifiers each for linearly separable and non-separable data and plotted the decision boundaries accordingly. The performance metrics for each of the classes