CS669 Assignment 1

Submitted by – Prachi Sharma

Branch – MSc. Applied Mathematics

Roll No. – V21078

Date – 10/10/2022

Q1. **You are going to perform speech activity detection (SAD.) Given a sequence of signal frames, classify each frame as speech or non-speech. Two types of 1-D features are provided: short-time energy, and Mellterbank energy. Which of these features are better at correctly detecting speech? Plot ROC curves to justify your choice.**

- **You can use a simple unimodal Gaussian to estimate the distribution of the features.**
- **Use sample mean and sample variance as parameters of the Gaussian.**
- **The ground truth _les are provided with 1 meaning speech and 0 meaning non-speech.**
- **Use Segment 2 for estimating the model and segment 3 for testing (ie ROC curves will be computed on Segment 3.)**

Expected outputs
Plot of ROC curves for each feature used.

Link to code notebook–

https://colab.research.google.com/drive/1FdHcotKWhGl7mShmhw3tDsmdi71Ht1mC?usp=sharing

Note: Please copy and paste the URL to open Colab Notebook.

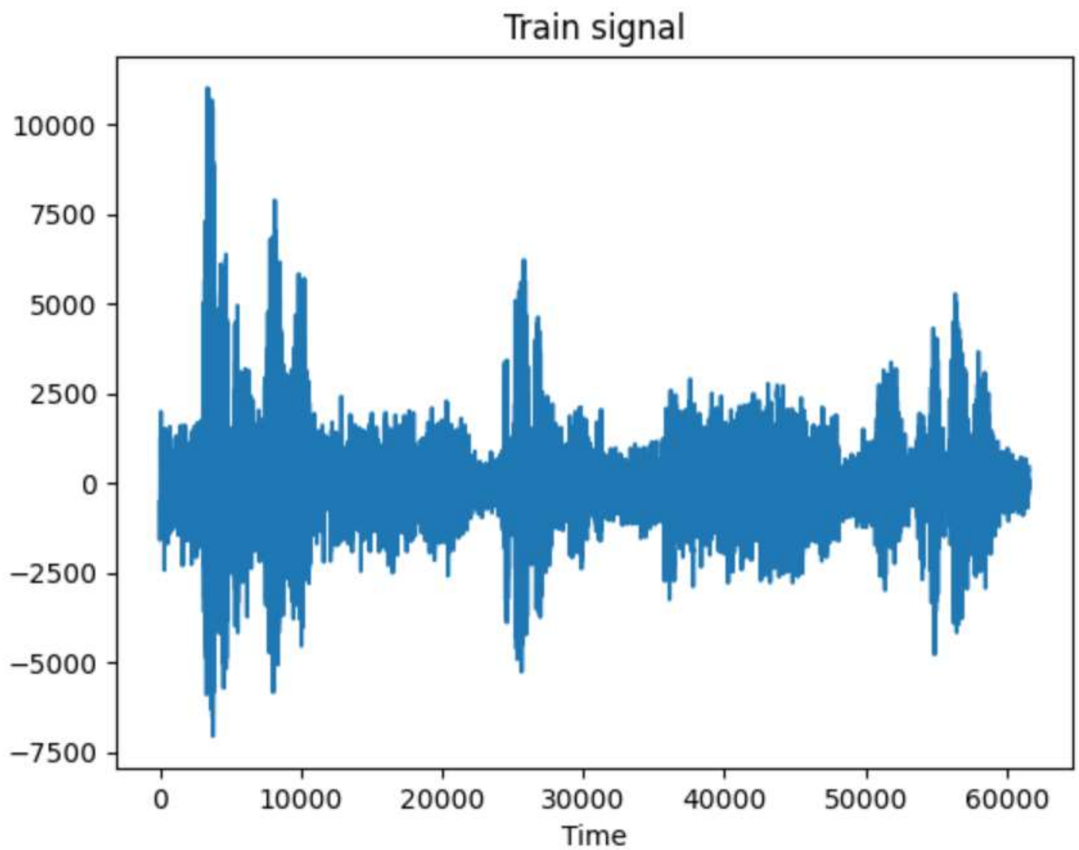
Observations and Results –

Since we are using two different types of energy data for the same Audio data. We have implemented two separate Unimodal Gaussian distributions to fit our training data and test on the second audio(Segment 3)'s both types of energy.

The two types of energy data available –

1. Short time energy (STE)
2. Mel Filter bank Energy (MEL)

Data Visualisation –



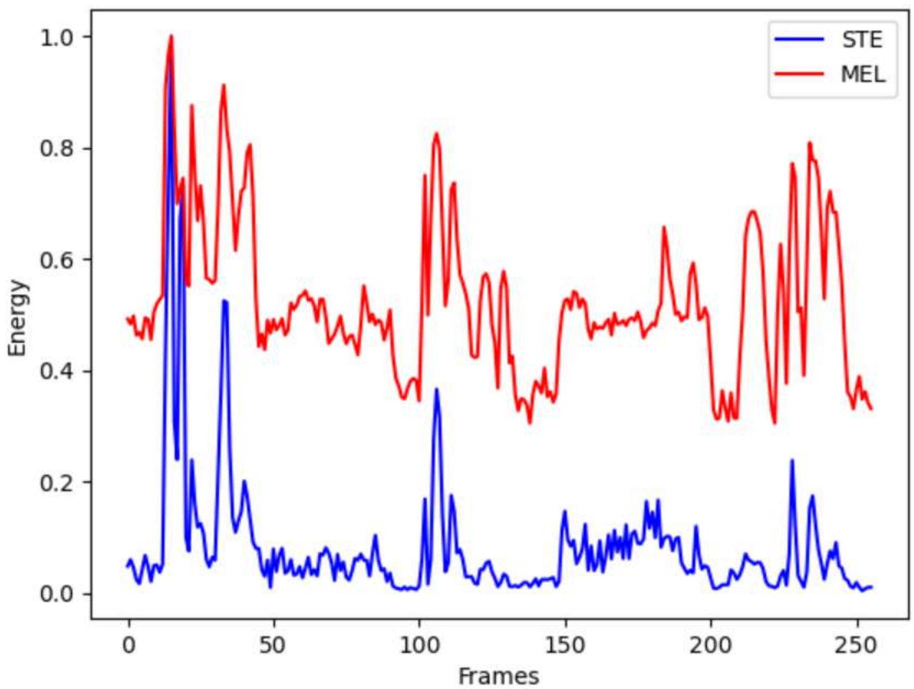
Above is the Segment 2 (Training data) Audio signal's Amplitude vs Time graph.

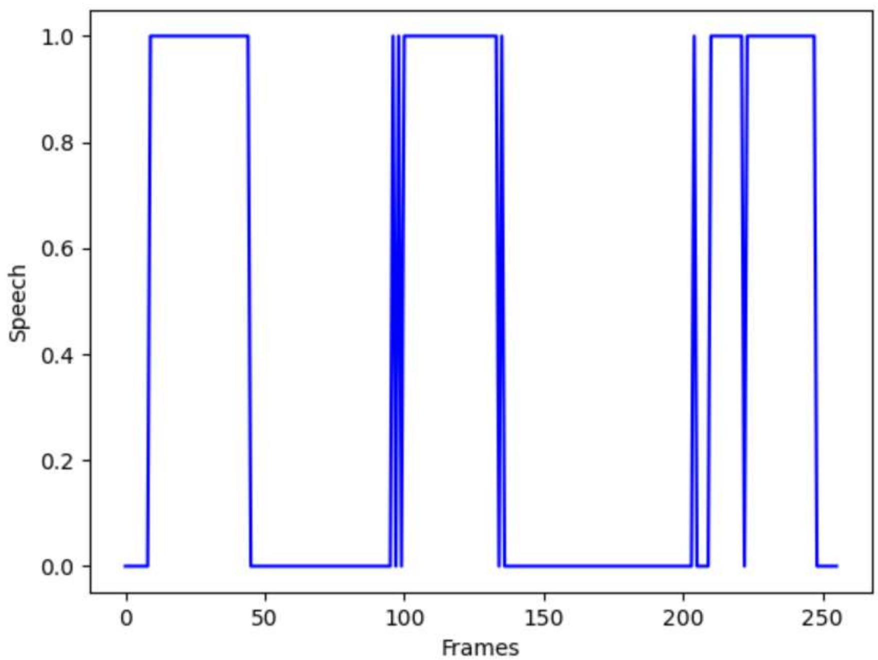
It can be observed that the data has **Sample rate = 8000.**

As mentioned in the question this audio signal is transformed to appropriate energy signals using Transformations.

STE Training data vs MEL Training data –

By Plotting the normalized values of energy for both cases -



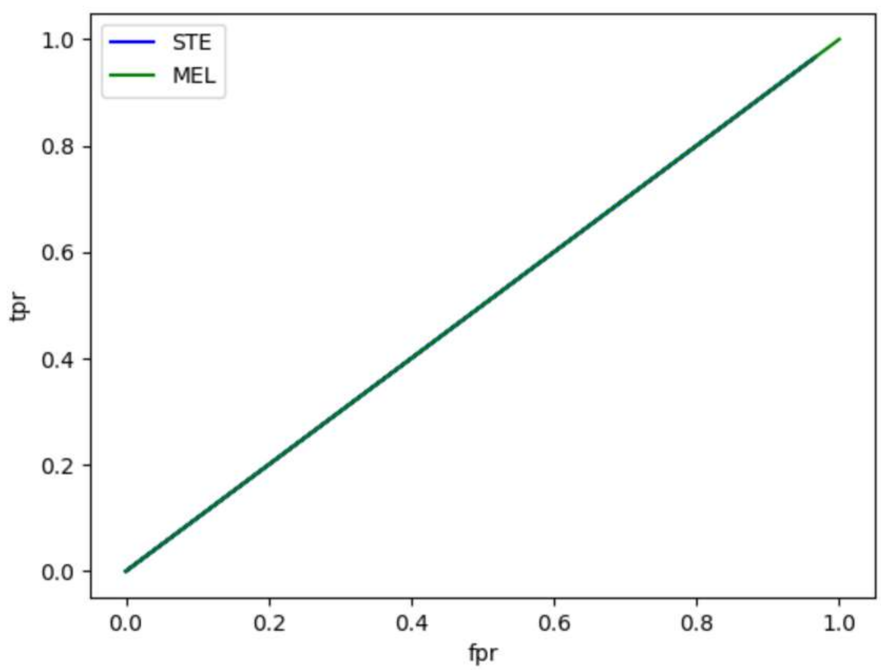


The Output data depicts if the binary value is 1 – it’s speech. If 0 – Non speech.

We have selected thresholds to be a uniform distribution with 10000 values. The output from the classifier gives us the probability for each instance of test data with a probability of the value being speech. This probability is derived from the gaussian distribution created by the sample mean and variance of the train data.

Plotting the ROC curve (by comparing output probability with the threshold value and assigning a class based on the fact that if the Probability (xi | Speech)> threshold) for both the energy band we need to observe the roc curve with the higher TPR values. Since TPR represents the actual true data predicted by our model.

Ideally, the curve should have different shapes. However, by our implementation, we observed the following overlapping ROC curve for both cases.



This however depicts that both cases are equally efficient. This may not be a correct estimate of the cases and can be possible because of some code error.

Q2. **Develop a Bayes classifier with Gaussian class conditional densities to classify two datasets, each having 3 classes. The first dataset is linearly separable, and the second is not. Use random 50% of data for training and 50% for testing. Build the following classifiers C1-C4:**

- **C1: Covariance for all classes is I_2. Use the average of the sample variances for all dimensions, for all classes, from the training data as _2.**
- **C2: Full but equal covariance for all classes, _. Use the average of the sample covariance matrix from all classes in the train data as _.**
- **C3: Diagonal covariance matrix, distinct for each class. Use variances from the sample covariance matrix for each class.**
- **C4: Full covariance matrix, distinct for each class. Use the sample covariance matrix for each class.**

Expected outputs

(a) Summarize the classifier performance as in Table 1. Use separate tables for linear and non-linear data.

(b) For each classifier, and for each dataset, plot the decision regions with class data in different colors. You will thus have 8 plots. A sample plot is shown in Figure 1.

Link to collab code –
https://colab.research.google.com/drive/1k5Cf8pykUPF75Ooup0mRfftKPriJnkmJ?usp=sharin g

Note: Please copy and paste the URL to open Colab Notebook.

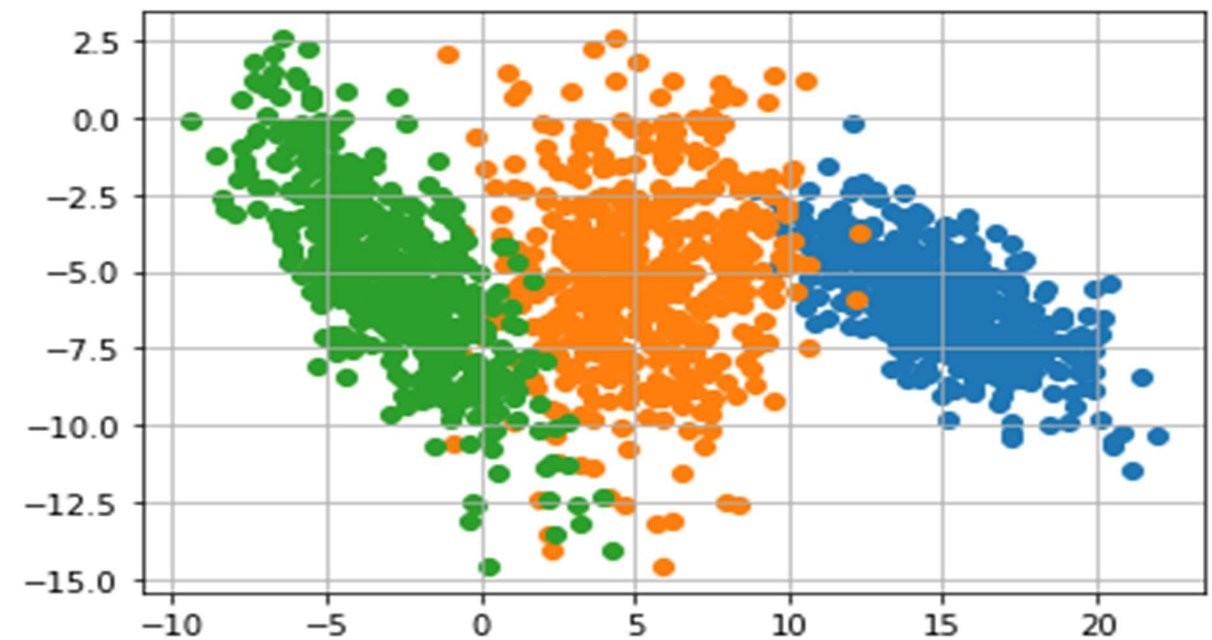
Data visualization and preprocessing –

Since for both the Linearly separable and Non linearly separable data. We have 3 separate files with two variables representing the 3 classes for each case.

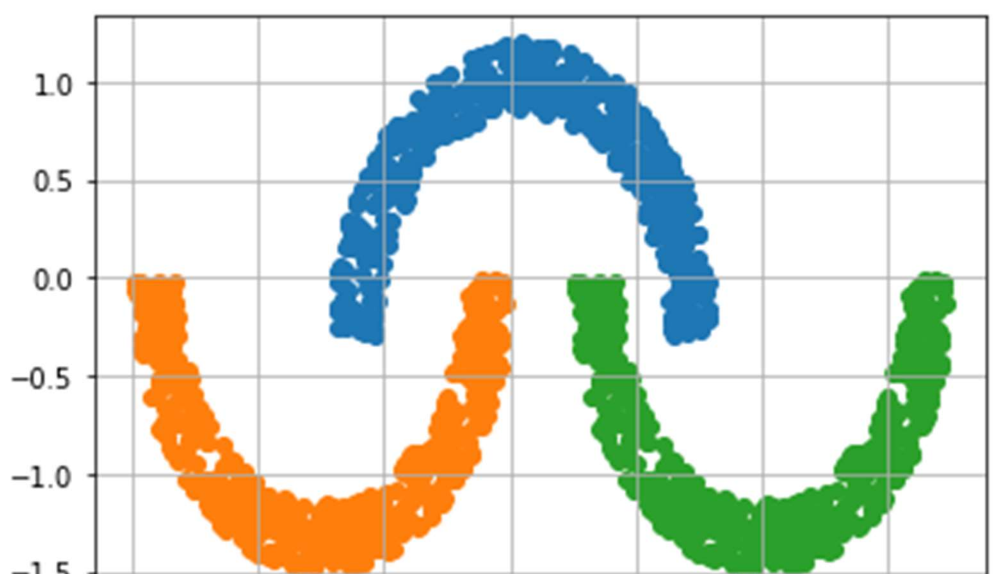
We have combined the 3 files for the linearly separable data together by adding a label of class to the three files as (1,2,3).

Hence our data looks like –
(X1, X2) , Y

The same has been implemented for the non-linearly separable case.

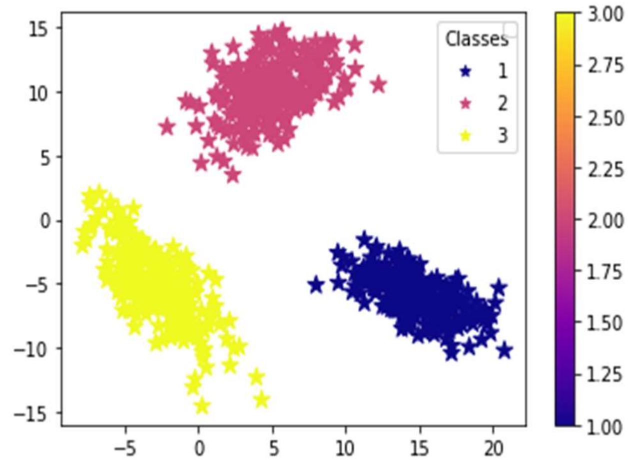


Three classes of the Linearly seperable data.
Total data points – 1500
Total variable – 2
Output Variable(class) – 1

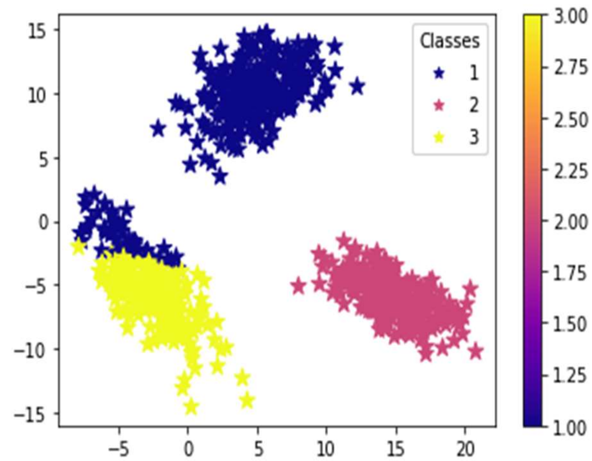


The three classes of the non-linearly separable data.
The data has further been split into 50% for training and 50% for testing randomly.

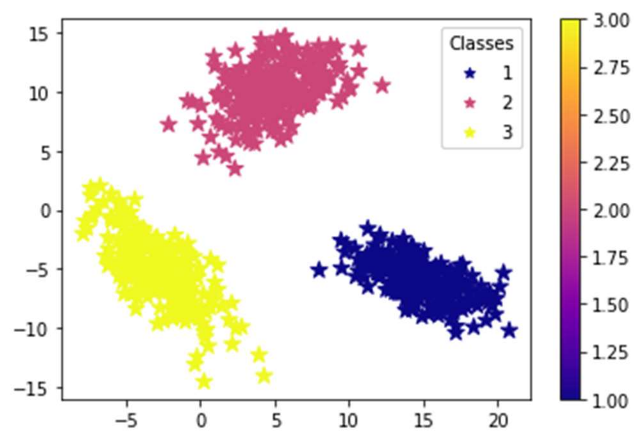
We have built 4 classifiers and each classifier would evaluate both type of data hence giving us 8 Plots.



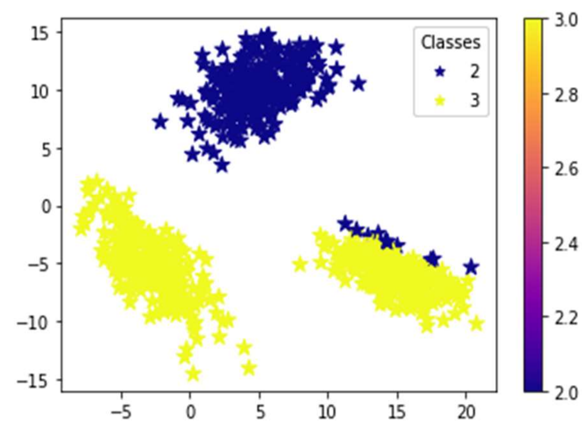
Classifier 1



Classifier 2

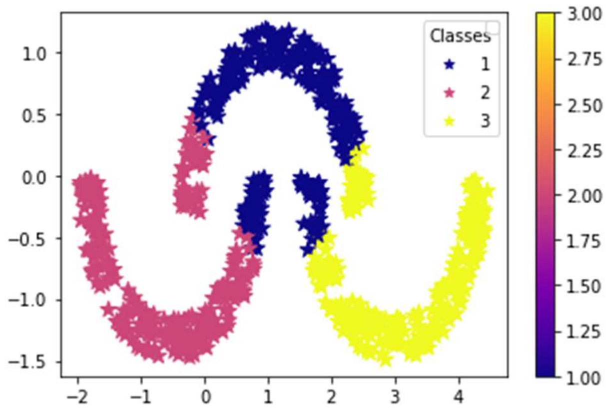


Classifier 3

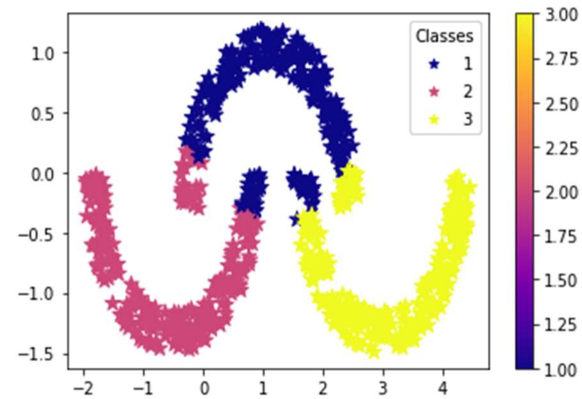


Classifier 4

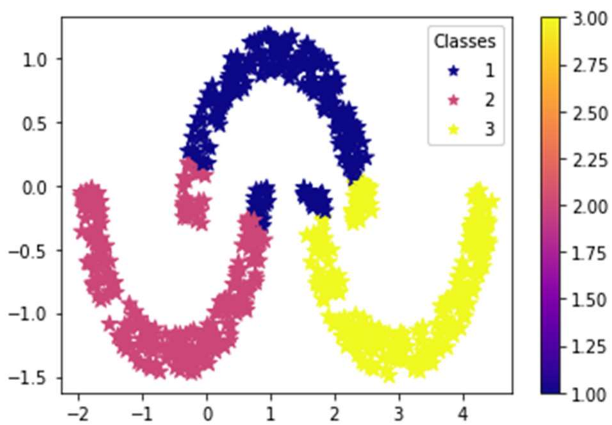
```
    Result Table for Linearly seperable Data
    ['Accuracy', 'Precision', 'Recall', 'F1 Score']
C1  [[1.         1.         1.         1.        ]
C2   [0.26533333 0.26533333 0.26533333 0.26533333]
C3   [1.         1.         1.         1.        ]
C4   [0.66533333 0.66533333 0.66533333 0.66533333]]
```



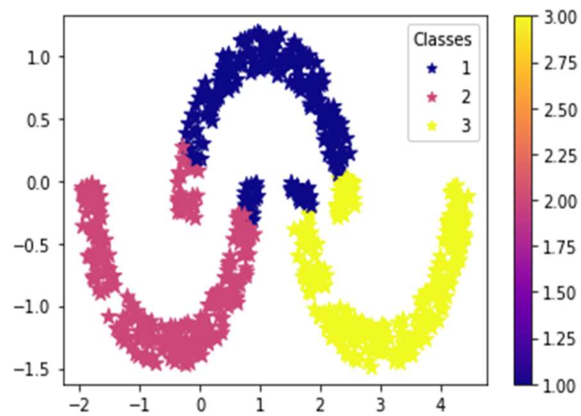
Classifier 1



Classifier 2



Classifier 3



Classifier 4

```
     Result Table for Non Linearly seperable Data
     ['Accuracy', 'Precision', 'Recall', 'F1 Score']
C1   [[0.808      0.808      0.808      0.808     ]
C2    [0.87466667 0.87466667 0.87466667 0.87466667]
C3    [0.87866667 0.87866667 0.87866667 0.87866667]
C4    [0.87733333 0.87733333 0.87733333 0.87733333]]
```

Inferences drawn –

- For linearly separable data, 1^{st} and 3^{rd} classifiers are performing the best. It is also represented in its accuracy, precision, recall, and F-score. This means it is having diagonal covariance matrix, i.e. features are uncorrelated.
- For non-linearly separable data, classifiers 3^{rd} and 4^{th} are giving better performance than the rest. Since it's non-linearly separable data, we are not getting perfect classification. Also, there exists some covariance in the features of data.