

## IML Assignment Report – Group 6

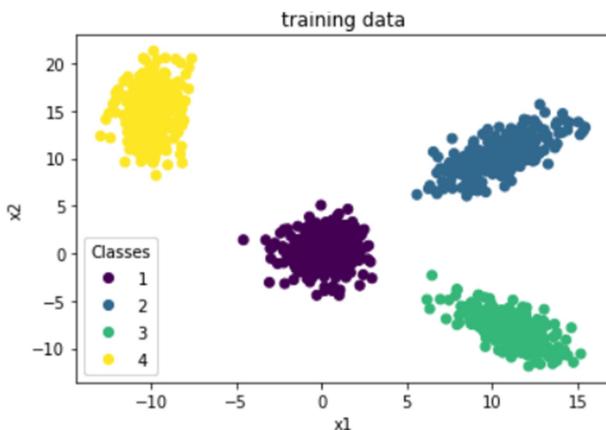
Group 6 consisted of two members- Jovian D’Souza (18CSE1012) and Sohail Hodarkar (18CSE1030). We were asked to build the following classifiers for 2-dimensional linearly separable, overlapping, and non-linearly separable datasets:

- Bayes Classifier
- Perceptron based Classifier
- SVM based Classifier

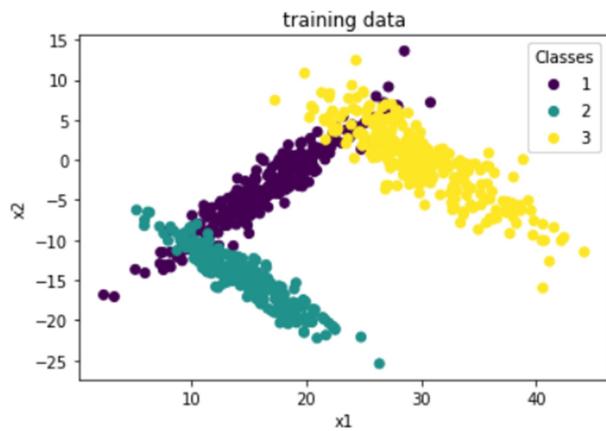
Our study consisted of several aspects, each of which will be covered in subsequent sections.

- **Plot of Training Data:**

Visualization of **Linearly Separable** Training Data:

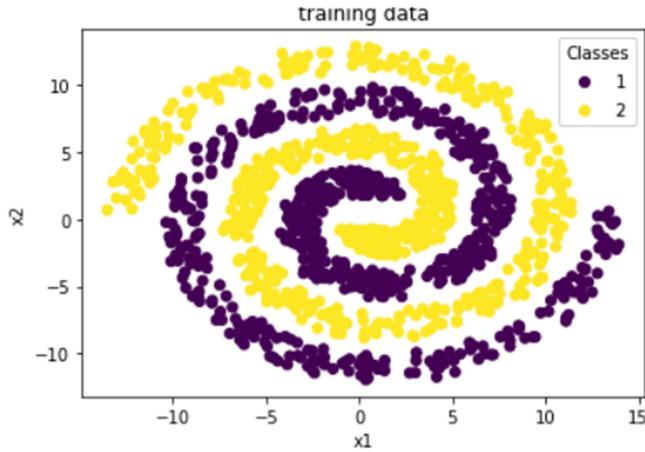


Visualization of **Overlapping** Training Data:



PTO...

Visualization of **Non-Linearly Separable** Training Data:



- **Classification Accuracy, Precision, Recall, and F-measure for each classifier:**

Performance of **Bayes Classifier** can be summarized as below:

#### **Linearly Separable Data:**

```
Mean Accuracy: 1.00
Mean Precision: 1.00
Mean Recall: 1.00
Mean F-measure: 1.00

Recall for class 1: 1.00
Recall for class 2: 1.00
Recall for class 3: 1.00
Recall for class 4: 1.00

F-measure for class 1: 1.00
F-measure for class 2: 1.00
F-measure for class 3: 1.00
F-measure for class 4: 1.00
```

#### **Overlapping Data:**

```
Mean Accuracy: 0.68
Mean Precision: 0.78
Mean Recall: 0.68
Mean F-measure: 0.68

Recall for class 1: 0.78
Recall for class 2: 1.00
Recall for class 3: 0.27

F-measure for class 1: 0.62
F-measure for class 2: 0.90
F-measure for class 3: 0.43
```

**PTO...**

### **Non-Linearly Separable Data:**

```
Mean Accuracy: 0.56
Mean Precision: 0.77
Mean Recall: 0.56
Mean F-measure: 0.56

Recall for class 1: 1.00
Recall for class 2: 0.12

F-measure for class 1: 0.70
F-measure for class 2: 0.22
```

Performance of **Perceptron based Classifier** can be summarized as below:

### **Linearly Separable Data:**

```
Mean Accuracy: 0.96
Mean Precision: 0.96
Mean Recall: 0.96
Mean F-measure: 0.96

Recall for class 1: 1.00
Recall for class 2: 1.00
Recall for class 3: 0.96
Recall for class 4: 0.87

F-measure for class 1: 0.92
F-measure for class 2: 1.00
F-measure for class 3: 0.98
F-measure for class 4: 0.93
```

### **Overlapping Data:**

```
Mean Accuracy: 0.88
Mean Precision: 0.88
Mean Recall: 0.88
Mean F-measure: 0.88

Recall for class 1: 0.82
Recall for class 2: 0.88
Recall for class 3: 0.94

F-measure for class 1: 0.82
F-measure for class 2: 0.88
F-measure for class 3: 0.93
```

### Non-Linearly Separable Data:

Mean Accuracy: 0.54  
Mean Precision: 0.54  
Mean Recall: 0.54  
Mean F-measure: 0.53

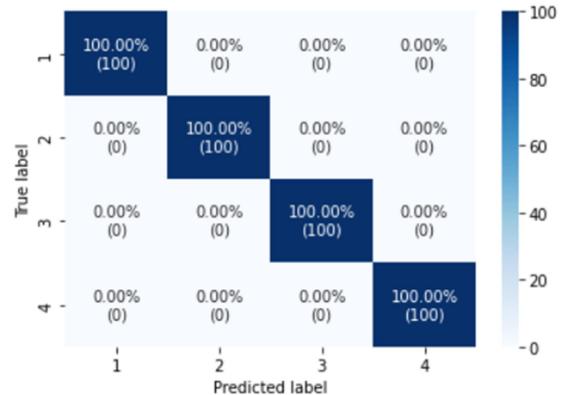
Recall for class 1: 0.42  
Recall for class 2: 0.67

F-measure for class 1: 0.47  
F-measure for class 2: 0.59

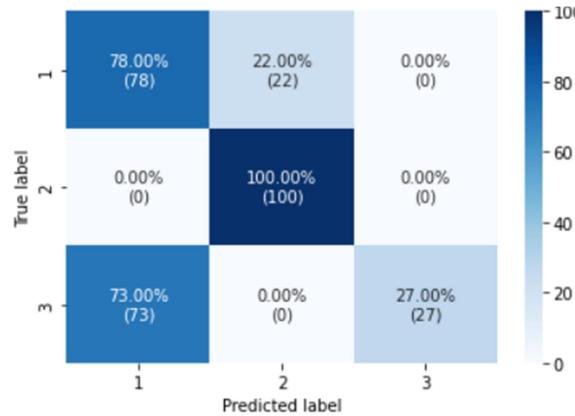
- **Confusion Matrix for test data performance:**

Confusion Matrix for Bayes Classifier is as follows:

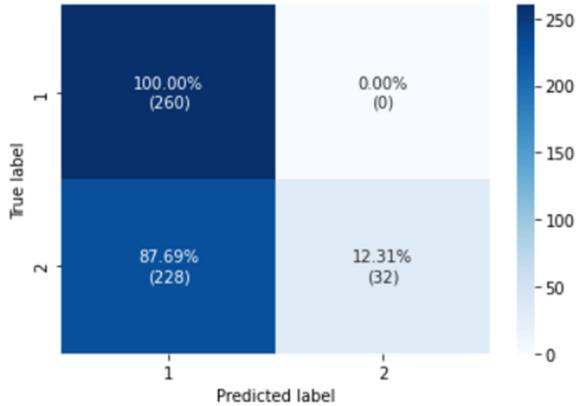
### Linearly Separable Data:



### Overlapping Data:

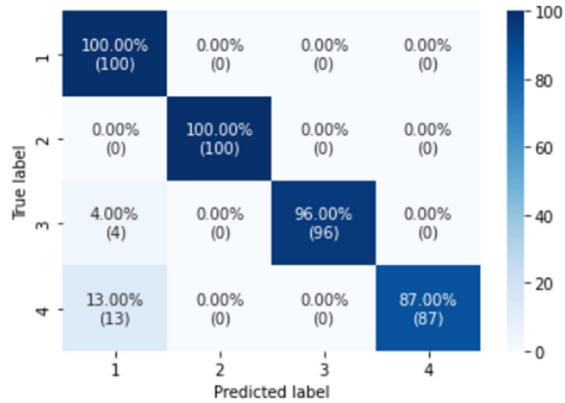


### Non-Linearly Separable Data:

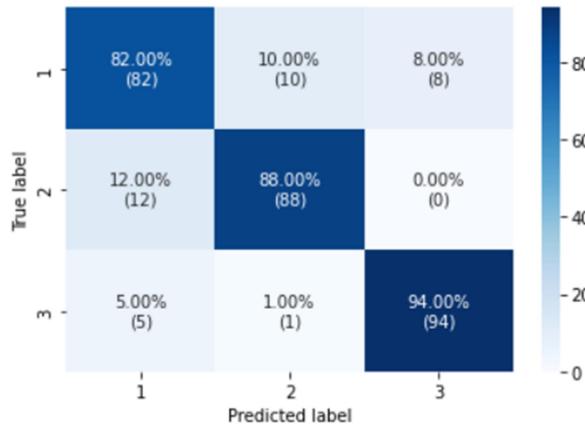


Confusion Matrix for Perceptron based Classifier is as follows:

### Linearly Separable Data:

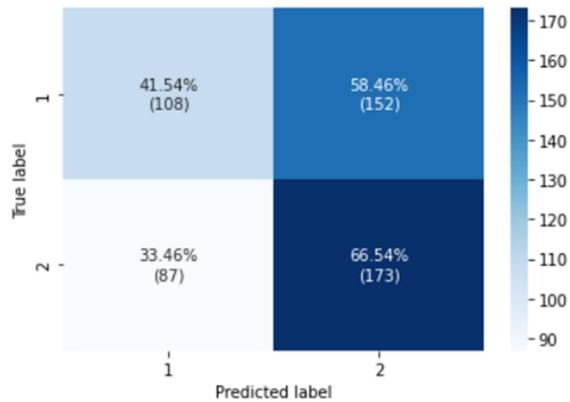


### Overlapping Data:



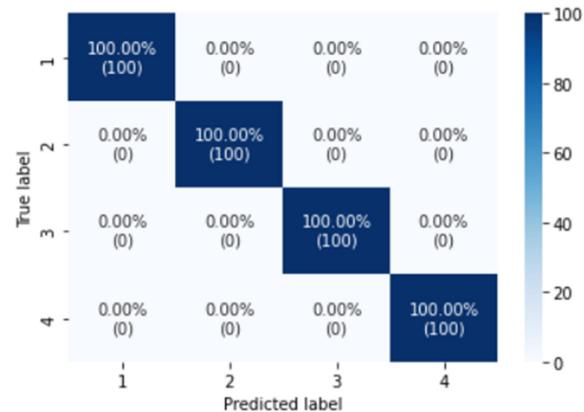
PTO...

### Non-Linearly Separable Data:

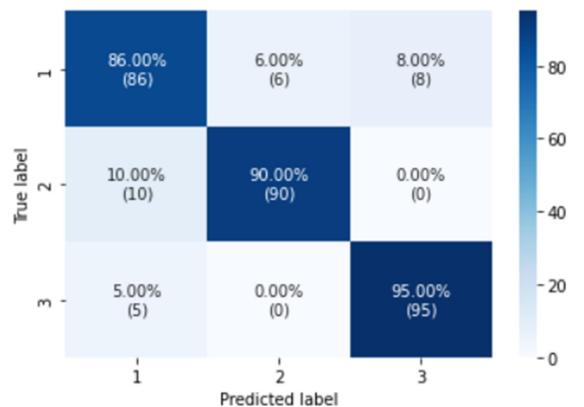


Confusion Matrix for **SVM based Classifier** is as follows:

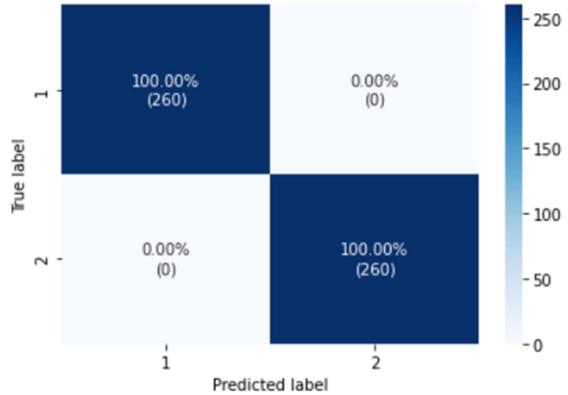
### Linearly Separable Data:



### Overlapping Data:



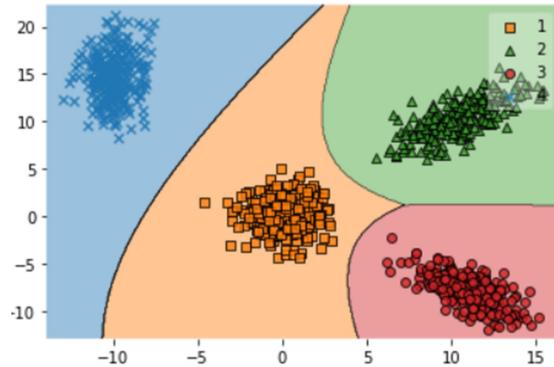
### Non-Linearly Separable Data:



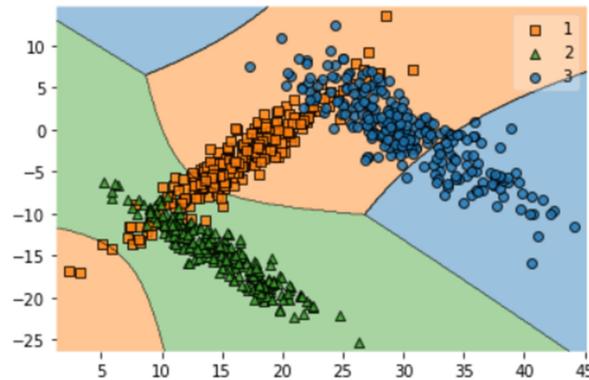
- **Decision Region for all classes together with training data superimposed:**

The decision region for **Bayes Classifier** with training data superimposed is as follows:

### Linearly Separable Data:

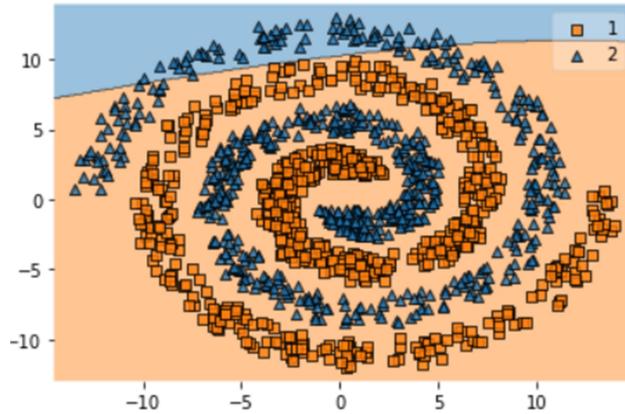


### Overlapping Data:



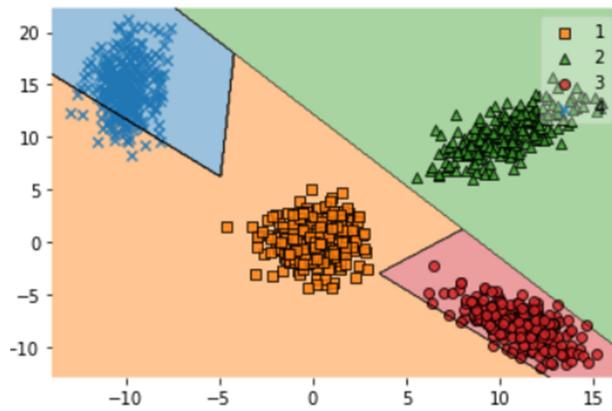
PTO...

### Non-Linearly Separable Data:

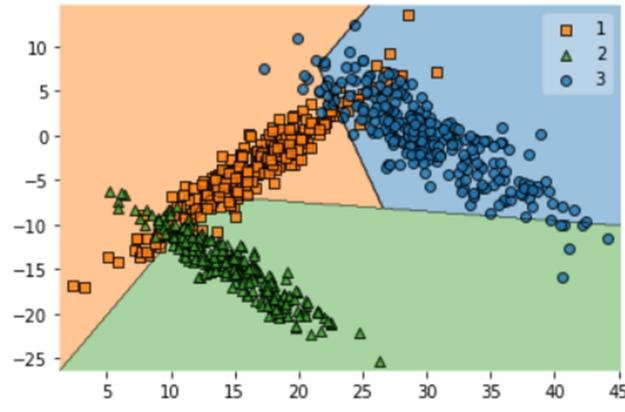


The decision region for **Perceptron based Classifier** with training data superimposed is as follows:

### Linearly Separable Data:

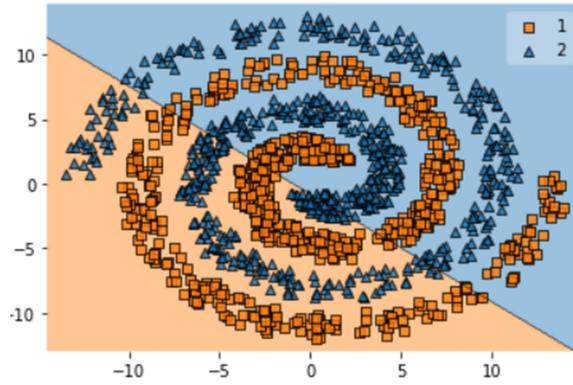


### Overlapping Data:



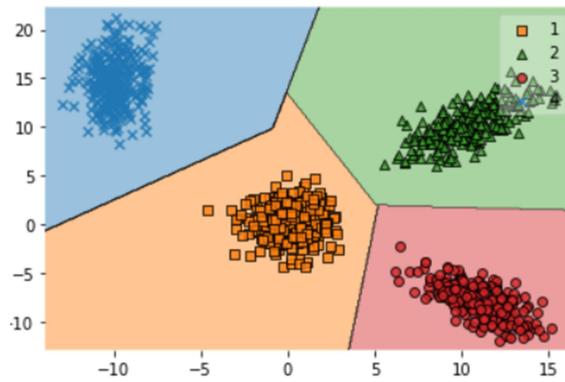
PTO...

### **Non-Linearly Separable Data:**

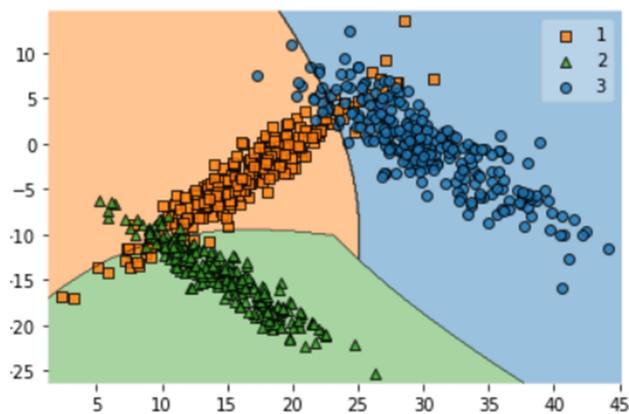


The decision region for **SVM based Classifier** with training data superimposed is as follows:

### **Linearly Separable Data:**

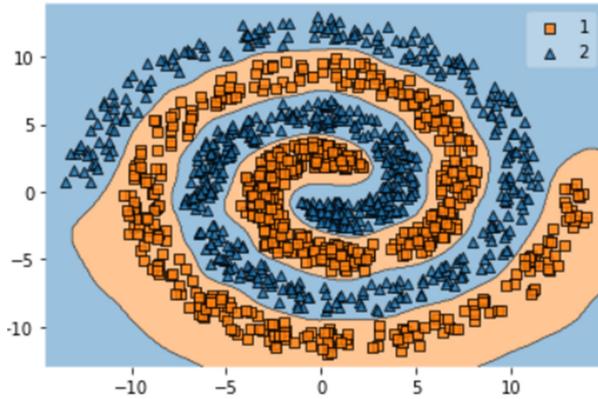


### **Overlapping Data:**



PTO...

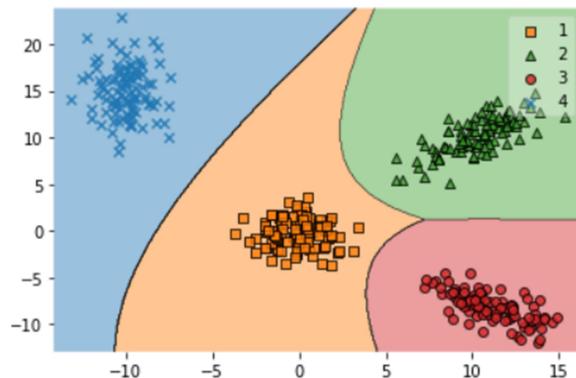
### Non-Linearly Separable Data:



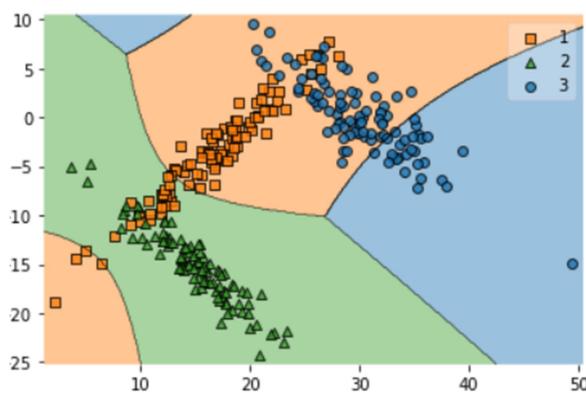
- Decision Region for all classes together with test data superimposed:

The decision region for **Bayes Classifier** with test data superimposed is as follows:

### Linearly Separable Data:

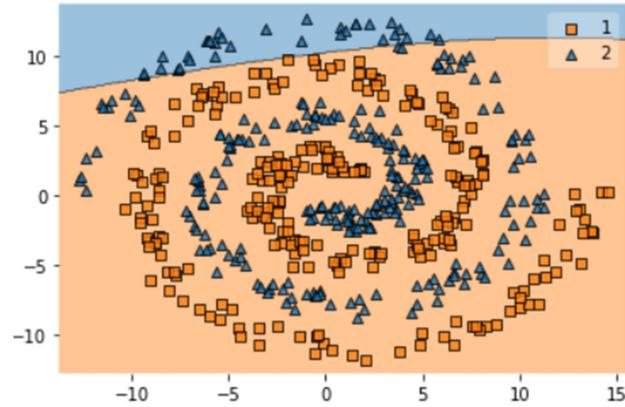


### Overlapping Data:



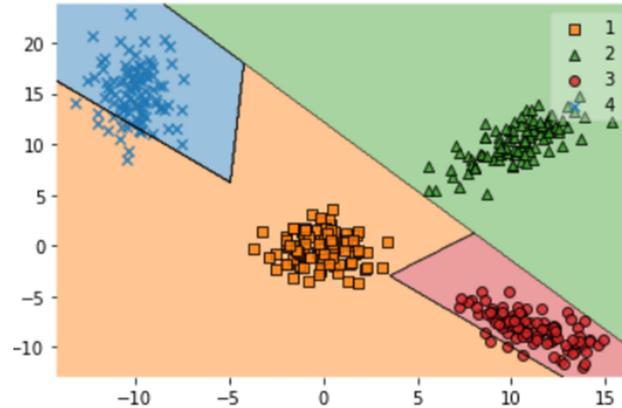
PTO...

### Non-Linearly Separable Data:

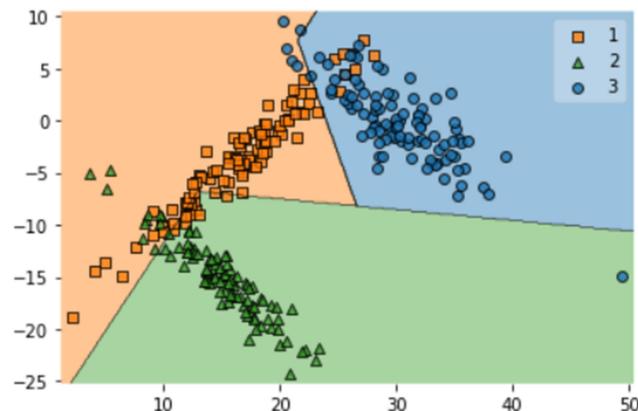


The decision region for **Perceptron based Classifier** with test data superimposed is as follows:

### Linearly Separable Data:

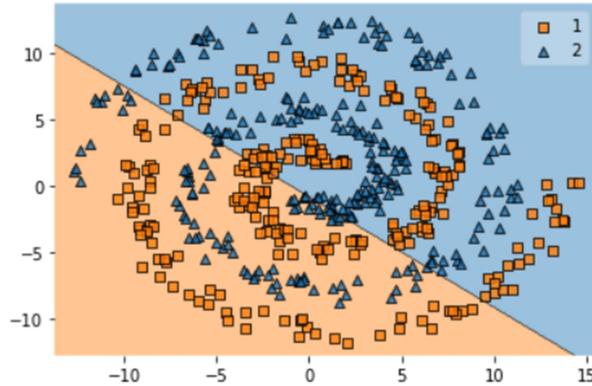


### Overlapping Data:



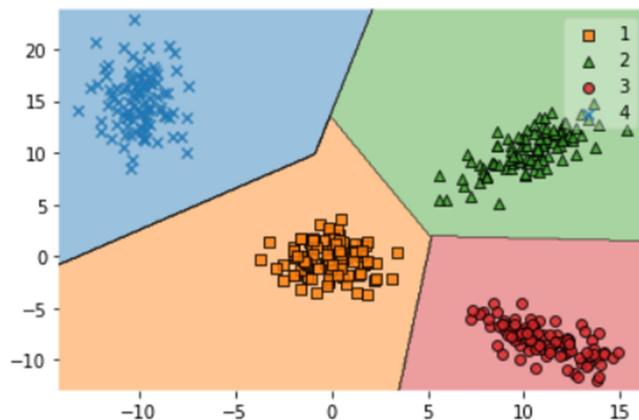
PTO...

### Non-Linearly Separable Data:

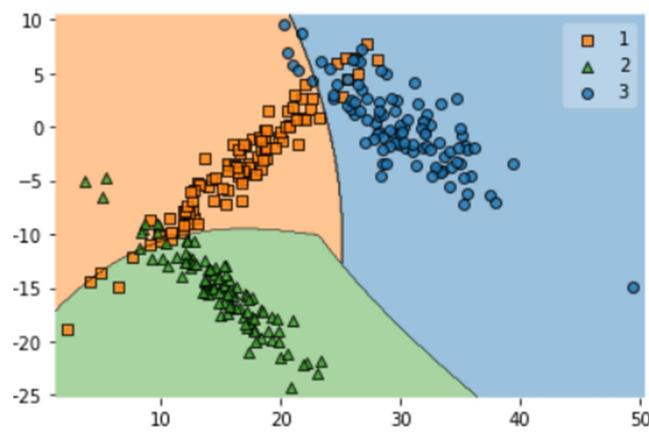


The decision region for **SVM based Classifier** with test data superimposed is as follows:

### Linearly Separable Data:

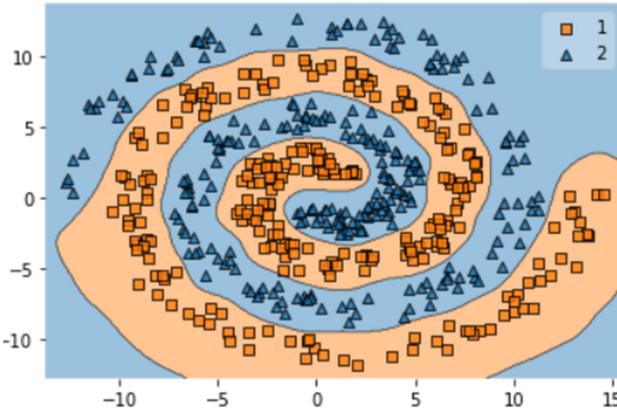


### Overlapping Data:



PTO...

### Non-Linearly Separable Data:



## Observations

It is interesting to note the nature of the decision surface for each classifier. The observations about the same may be summarized as follows:

### Bayes Classifier:

The Bayes Classifier is seen to generate the following decision surfaces:

- a. **Linearly Separable Data:** Piecewise Linear decision boundary
- b. **Overlapping Data:** Piecewise Linear decision boundary
- c. **Non-Linearly Separable Data:** Piecewise Linear decision boundary

### Perceptron based Classifier:

The Perceptron based classifier is seen to generate the following decision surfaces:

- a. **Linearly Separable Data:** Linear decision boundary
- b. **Overlapping Data:** Linear decision boundary
- c. **Non-Linearly Separable Data:** Linear decision boundary

### SVM based Classifier:

The SVM based classifier is seen to generate the following decision surfaces, depending on the kernel being used:

- a. **Linearly Separable Data:** Linear decision boundary
- b. **Overlapping Data:** 3<sup>rd</sup> degree polynomial decision boundary
- c. **Non-Linearly Separable Data:** Radial non-linear decision boundary