**Dry Beans Classifier – NN – Task 1**

**TeamID: 11**

**Team Members**

|  |  |  |
| --- | --- | --- |
| Name | ID | Department |
| أحمد حسن حسن خطاب | *20201700033* | Computer Science(CS). |
| انس سعيد صبري | *20201700159* | Computer Science(CS). |
| عبدالرحمن هاني خيري الدين | *20201700488* | Computer Science(CS). |
| أحمد هاني حامد محمد | *20191700875* | Computer Science(CS). |
| عمر رضا عبدالدايم عبدالعزيز | *20201701267* | Computer Science(CS). |

1. **Single Layer Perceptron**

**Combination#1**

* **Inputs:**

**Selected features: Area , Perimeter.**

**Selected classes: BOMBAY , CALI (C1&C2).**

**Learning rate: 0.2.**

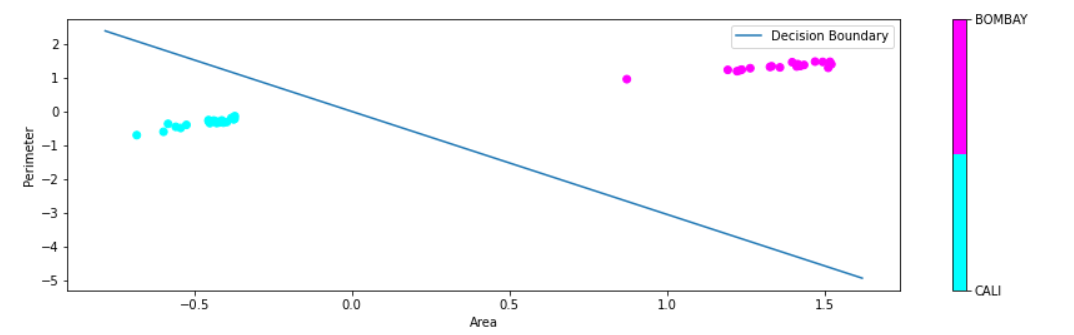
**Number of epochs: 10.**

**MSE threshold: 3.**

**Use Bias ?: Yes.**

**Chosen algorithm: perceptron.**

* **Testing data visualization:**

****

* **Model evaluation:**

**Accuracy: 100%.**

**Confusion Matrix:**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Predicted** | | |
| **Actual** |  | **Positive** | **Negative** |
| **Positive** | **20** | **0** |
| **Negative** | **0** | **20** |

* **Analysis:**
* **The Inputs had make the algorithm plotting a good decision boundary causing no overfitting.**
* **This means that Area & Perimeter are good features to classify BOMBAY & CALI dry beans.**

**Combination#2**

* **Inputs:**

**Selected features: Area , Perimeter.**

**Selected classes: CALI , SIRA (C2&C3).**

**Learning rate: 0.2.**

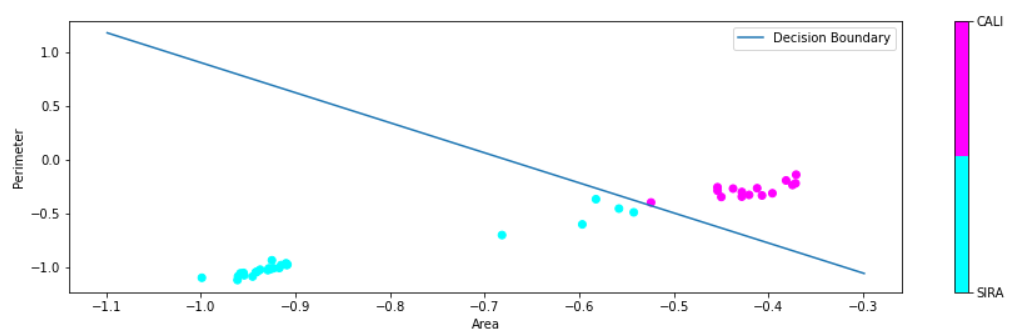
**Number of epochs: 10.**

**MSE threshold: 3.**

**Use Bias ?: Yes.**

**Chosen algorithm: perceptron.**

* **Testing data visualization:**

****

* **Model evaluation:**

**Accuracy: 87.5%.**

**Confusion Matrix:**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Predicted** | | |
| **Actual** |  | **Positive** | **Negative** |
| **Positive** | **15** | **5** |
| **Negative** | **0** | **20** |

* **Analysis:**
* **The Inputs had make the model higher precision and lower recall than the previous model.**
* **This means that Area & Perimeter are not good features to classify between CALI&SIRA dry beans as it biases to SIRA (Negative class).**

**Combination#3**

* **Inputs:**

**Selected features: MajorAxisLength , MinorAxisLength.**

**Selected classes: CALI , SIRA (C2&C3).**

**Learning rate: 0.25.**

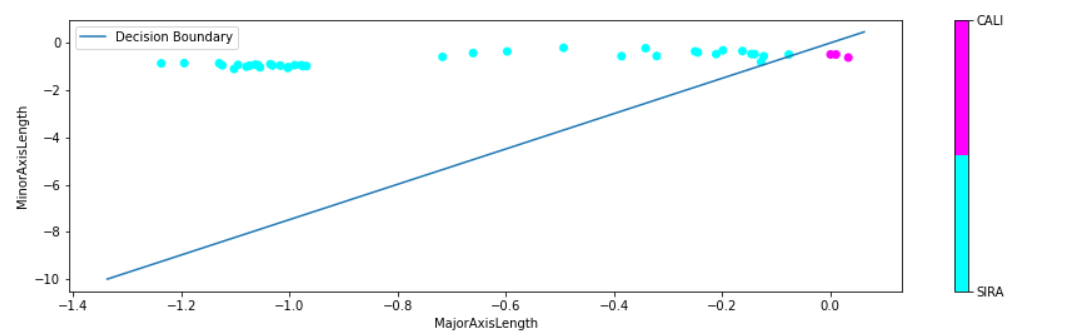
**Number of epochs: 10.**

**MSE threshold: 3.**

**Use Bias ?: No.**

**Chosen algorithm: perceptron.**

* **Testing data visualization:**

****

* **Model evaluation:**

**Accuracy: 57.5%.**

**Confusion Matrix:**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Predicted** | | |
| **Actual** |  | **Positive** | **Negative** |
| **Positive** | **3** | **17** |
| **Negative** | **0** | **20** |

* **Analysis:**
* **The Inputs had make the model more pessimistic than combination#2 model , which mean it had very high precision and very low recall.**
* **This means that MajorAxisLength & MinorAxisLength [with no bias] are not good features to classify between CALI&SIRA dry beans as it is so pessimistic model.**

**Combination#4**

* **Inputs:**

**Selected features: MajorAxisLength , MinorAxisLength.**

**Selected classes: CALI , SIRA (C2&C3).**

**Learning rate: 0.25.**

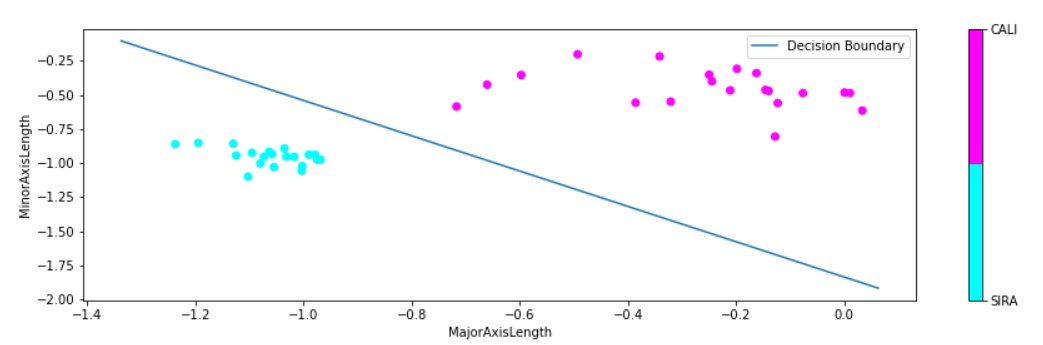
**Number of epochs: 10.**

**MSE threshold: 0.**

**Use Bias ?: Yes.**

**Chosen algorithm: perceptron.**

* **Testing data visualization:**

****

* **Model evaluation:**

**Accuracy: 100%.**

**Confusion Matrix:**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Predicted** | | |
| **Actual** |  | **Positive** | **Negative** |
| **Positive** | **20** | **0** |
| **Negative** | **0** | **20** |

* **Analysis:**
* **The Inputs had make the algorithm plotting a good decision boundary causing no overfitting.**
* **This means that MajorAxisLength & MinorAxisLength [with bias] are good features to classify CALI&SIRA dry beans.**

**Combination#5**

* **Inputs:**

**Selected features: Perimeter , roundnes.**

**Selected classes: BOMBAY , SIRA (C1&C3).**

**Learning rate: 0.25.**

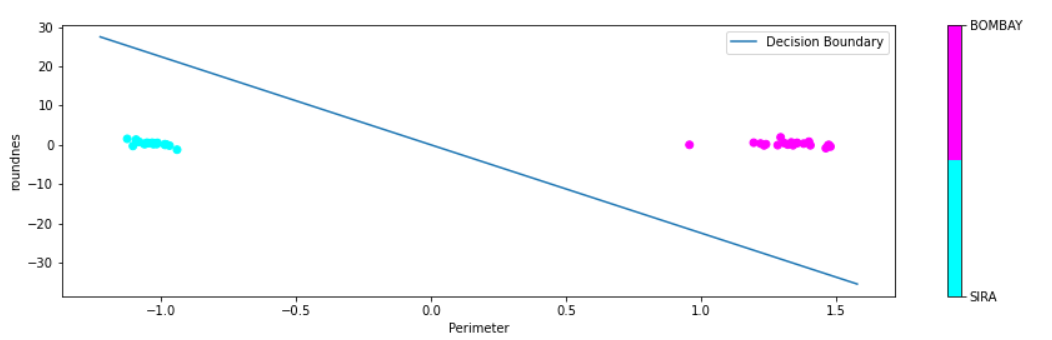
**Number of epochs: 10.**

**MSE threshold: 25.**

**Use Bias ?: No.**

**Chosen algorithm: perceptron.**

* **Testing data visualization:**

****

* **Model evaluation:**

**Accuracy: 100%.**

**Confusion Matrix:**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Predicted** | | |
| **Actual** |  | **Positive** | **Negative** |
| **Positive** | **20** | **0** |
| **Negative** | **0** | **20** |

* **Analysis:**
* **The Inputs had make the algorithm plotting a good decision boundary causing no overfitting.**
* **This means that Perimeter & roundnes are good features to classify CALI&SIRA dry beans.**

1. **Adaline**

**Combination#1**

* **Inputs:**

**Selected features: Area , Perimeter.**

**Selected classes: BOMBAY , CALI (C1&C2).**

**Learning rate: 0.25.**

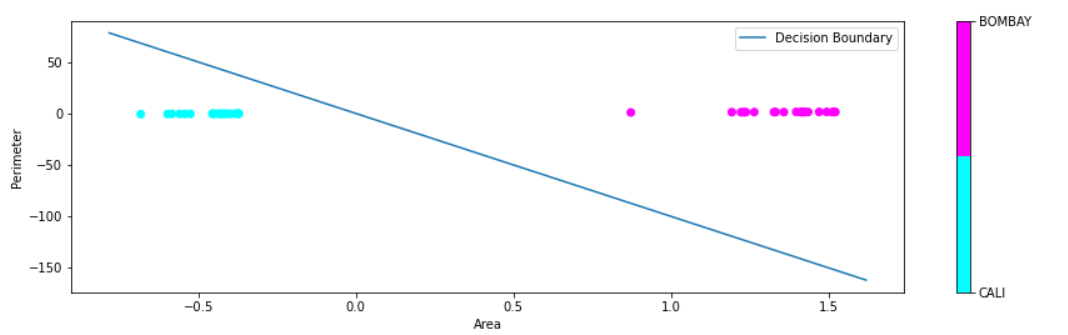
**Number of epochs: 10.**

**MSE threshold: 2.**

**Use Bias ?: No.**

**Chosen algorithm: Adaline.**

* **Testing data visualization:**

****

* **Model evaluation:**

**Accuracy: 100%.**

**Confusion Matrix:**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Predicted** | | |
| **Actual** |  | **Positive** | **Negative** |
| **Positive** | **20** | **0** |
| **Negative** | **0** | **20** |

* **Analysis:**
* **The Inputs had make the algorithm plotting a good decision boundary causing no overfitting.**
* **This means that Area & Perimeter are good features to classify BOMBAY&CALI dry beans.**

**Combination#2**

* **Inputs:**

**Selected features: Area , Perimeter.**

**Selected classes: CALI , SIRA (C2&C3).**

**Learning rate: 0.25.**

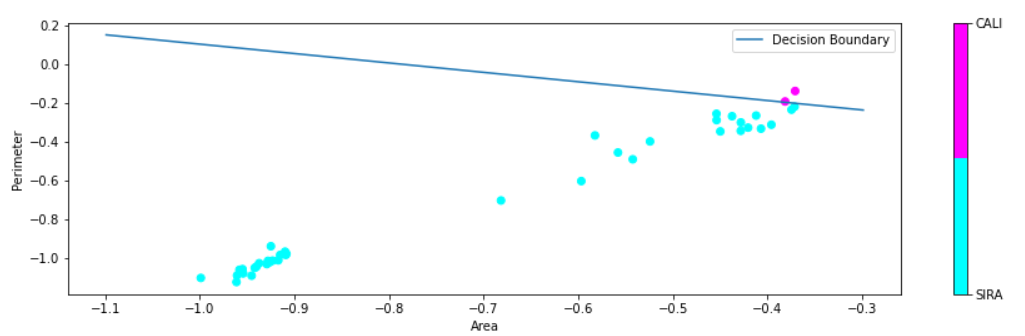
**Number of epochs: 10.**

**MSE threshold: 2.**

**Use Bias ?: Yes.**

**Chosen algorithm: Adaline.**

* **Testing data visualization:**

****

* **Model evaluation:**

**Accuracy: 55%.**

**Confusion Matrix:**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Predicted** | | |
| **Actual** |  | **Positive** | **Negative** |
| **Positive** | **2** | **18** |
| **Negative** | **0** | **20** |

* **Analysis:**
* **This model is about to be so pessimistic model as it biases to the Negative class (SIRA).**
* **Area & Perimeter are not a good features to classify CALI&SIRA dry beans.**

**Combination#3**

* **Inputs:**

**Selected features: MajorAxisLength , MinorAxisLength.**

**Selected classes: CALI , SIRA (C2&C3).**

**Learning rate: 0.25.**

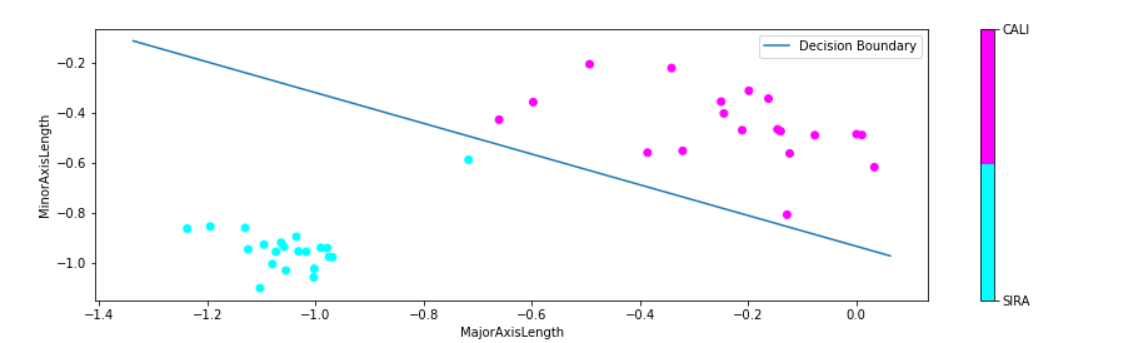
**Number of epochs: 10.**

**MSE threshold: 0.**

**Use Bias ?: Yes.**

**Chosen algorithm: Adaline.**

* **Testing data visualization:**

****

* **Model evaluation:**

**Accuracy: 97.5%.**

**Confusion Matrix:**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Predicted** | | |
| **Actual** |  | **Positive** | **Negative** |
| **Positive** | **19** | **1** |
| **Negative** | **0** | **20** |

* **Analysis:**
* **MajorAxisLength & MinorAxisLength are better than Area & Perimeter to classify CALI&SIRA dry beans.**

**Combination#4**

·       **Inputs:**

**Selected features: roundnes , Perimeter.**

**Selected classes: BOMBAY , SIRA (C1&C3).**

**Learning rate: 0.25.**

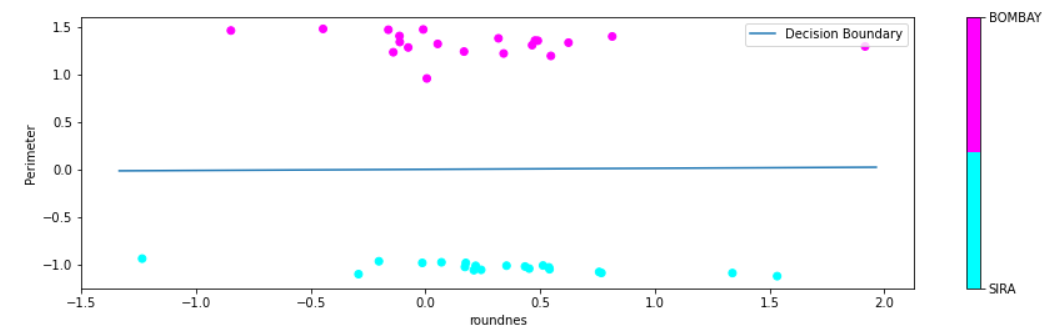
**Number of epochs: 10.**

**MSE threshold: 0.**

**Use Bias ?: No.**

**Chosen algorithm: Adaline.**

·       **Testing data visualization:**



·       **Model evaluation:**

**Accuracy: 100%.**

**Confusion Matrix:**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Predicted** | | |
| **Actual** |  | **Positive** | **Negative** |
| **Positive** | **20** | **0** |
| **Negative** | **0** | **20** |

·       **Analysis:**

* **The Inputs had make the algorithm plotting a good decision boundary causing no overfitting.**
* **This means that roundnes & Perimeter are good features to classify BOMBAY&SIRA dry beans.**

**Combination#5**

·       **Inputs:**

**Selected features: Area , MinorAxisLength.**

**Selected classes: BOMBAY , SIRA (C1&C3).**

**Learning rate: 0.25.**

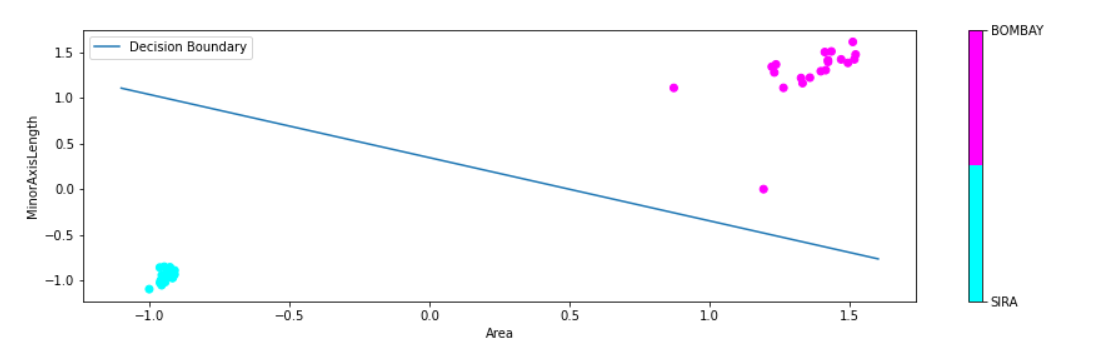
**Number of epochs: 10.**

**MSE threshold: 0.**

**Use Bias ?: Yes.**

**Chosen algorithm: Adaline.**

·       **Testing data visualization:**



·       **Model evaluation:**

**Accuracy: 100%.**

**Confusion Matrix:**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Predicted** | | |
| **Actual** |  | **Positive** | **Negative** |
| **Positive** | **20** | **0** |
| **Negative** | **0** | **20** |

·       **Analysis:**

* **The Inputs had make the algorithm plotting a good decision boundary causing no overfitting.**
* **This means that Area & MinorAxisLength are good features to classify BOMBAY&SIRA dry beans.**

**Conclusion**

* **Area & Perimeter** are the best features to classify between **BOMBAY[C1] & CALI[C2]** dry beans classes in both SLP and Adaline algorithms.
* **MajorAxisLength & MinorAxisLength** are the best features to classify between **CALI[C2] & SIRA[C3]** dry beans classes in both SLP and Adaline algorithms **provided** using the bias in training process but they give a better accuracy in SLP.
* **SLP** is more accurate model than Adaline in classification between **CALI[C2] & SIRA[C3]** using the same features and hyperparameters.
* **Any** combination of two features can classify between **BOMBAY[C1] & SIRA[C3]** dry beans classes in both SLP and Adaline algorithms.
* Using of the **bias** in training calculations make the model more accurate than ignoring the bias.
* In this binary classification task the **best** NN algorithm that gives better testing accuracy in general is **Single Layer Perceptron**.