

Single-Layer Perceptron (SLP) and ADALINE for Bird Species Classification

Task-1 Report

Team Members:

Zeina Shawkat Mohamed Nabil	2022170818	AI3
Menna Ali Thabet	2022170844	AI3
Haneen Akram	2022170812	AI3
Reem Ahmed Ismail	2022170815	AI3
Mohamed Ashraf Fathy	2022170919	AI3
Eslam Hesham Mohamed abdelazim	2022170802	AI3

Introduction

This report details a data analysis project focused on classifying bird species using a dataset with 150 records and features such as `gender`, `body_mass`, `beak_length`, `beak_depth`, and `fin_length`. The primary goal was to explore the dataset, visualize feature distributions, and implement Single-Layer Perceptron (SLP) and Adaline algorithms for classification. We also aimed to identify the features that contribute most to accurate classification.

Features Visualization

- **Class Distribution:** A pie chart (Figure 1) represents the proportion of males and females in the dataset. The gender distribution is even. An even gender distribution helps ensure that the classification models are not biased towards one gender, leading to more generalizable results. This balance suggests that gender may not be a strong discriminating feature on its own.

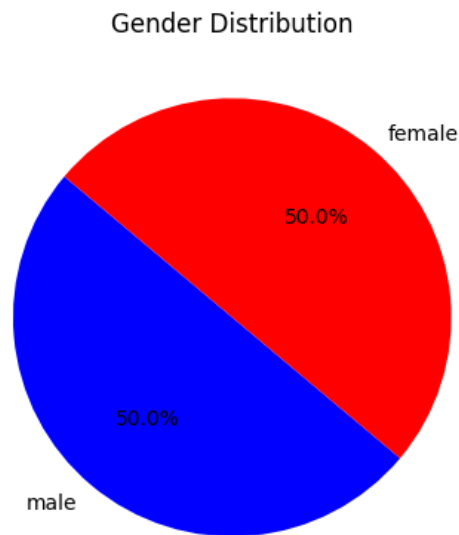


Figure 1: Gender Distribution Pie Chart

- **Body Mass Distribution:** A histogram (Figure 2) shows the distribution of `body_mass`. The distribution is right-skewed. This indicates that a majority of the birds have lower body mass, but there are some birds with significantly higher body masses, creating a tail on the right side of the distribution.

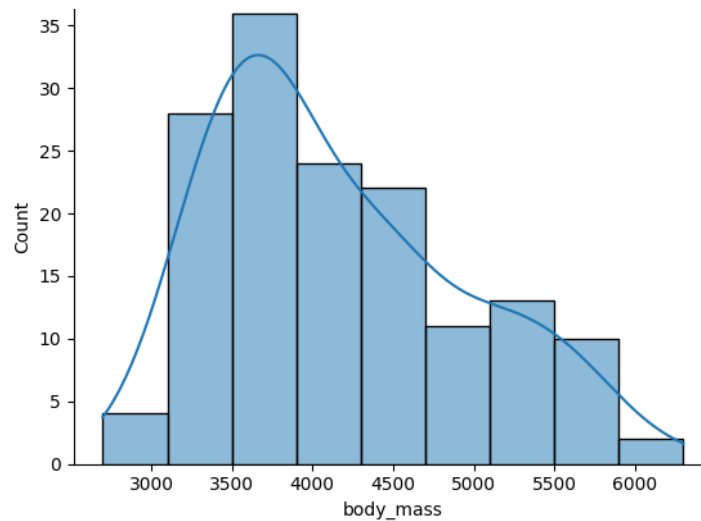


Figure 2: Body Mass Distribution Histogram

- **Beak Depth Distribution:** A histogram (Figure 3) displays the distribution of `beak_depth`. The distribution is left-skewed. This indicates that larger beak depths are less common than smaller beak depths.

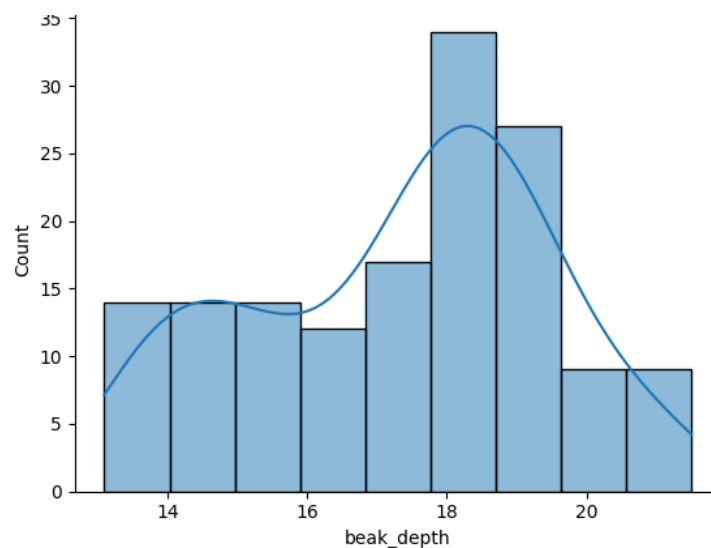


Figure 3: Beak Depth Distribution Histogram

- **Fin Length Distribution:** A histogram (Figure 4) displays the distribution of `fin_length`. It can be inferred that it shows the frequency of different fin lengths in the dataset. This distribution allowed for informed decisions about feature scaling and potential outlier handling.

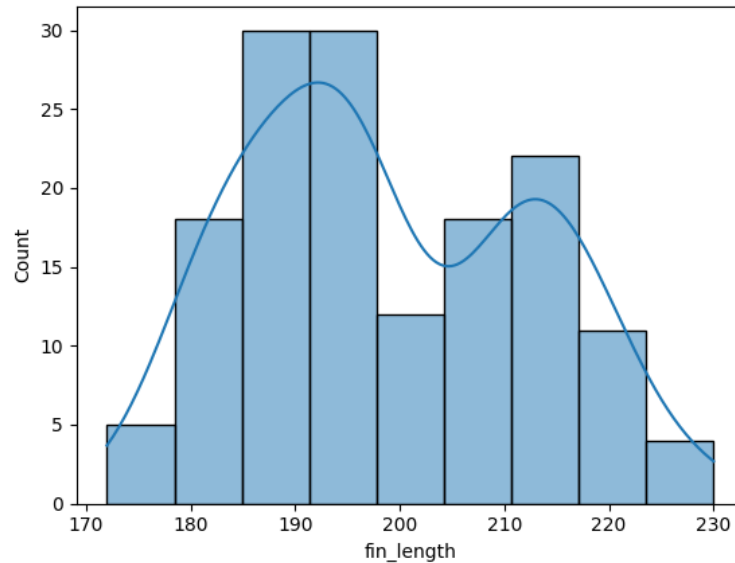


Figure 4: Fine Length Distribution Histogram

- **Bird Category Distribution:** A pie chart (Figure 5) shows the distribution of the different bird categories (A, B, C). The classes are evenly distributed. Specifically, each category represents 33.3% of the dataset. This indicates a balanced dataset concerning the target variable. A balanced class distribution is ideal for training classification models, as it prevents the model from being biased towards the majority class.

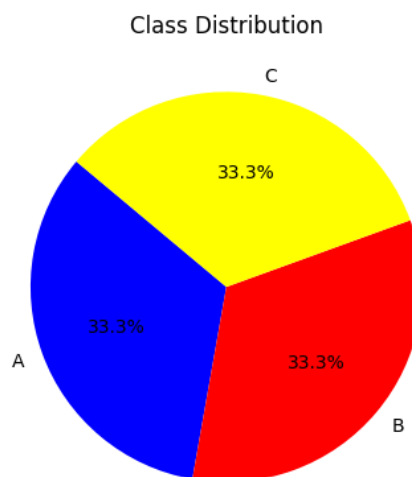


Figure 5: Class Distribution Pie Chart

- Box Plots of Features:** Box plots at (Figure 6) were generated for each feature (`gender`, `body_mass`, `beak_length`, `beak_depth`, `fin_length`, and `bird category`). Box plots visually represent the median, quartiles, and potential outliers in each feature's distribution. They help identify features with high variability or outliers, which might require special attention during preprocessing (e.g., outlier removal or winsorization).

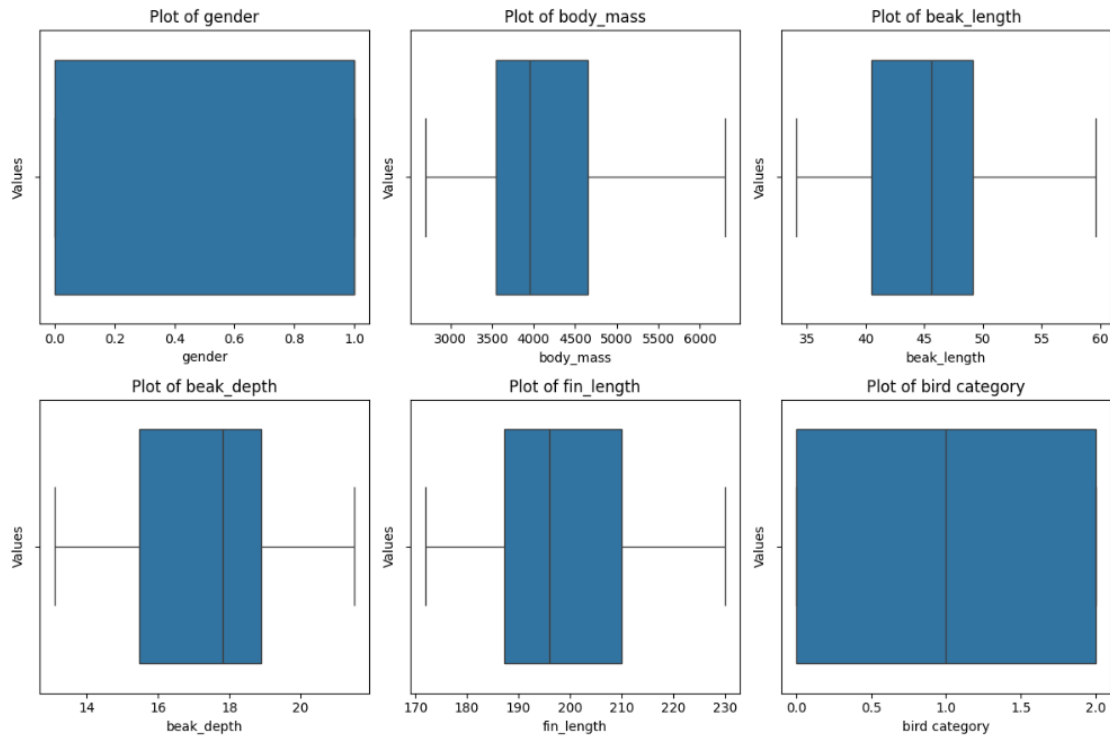


Figure 6: Box Plots Distribution Histogram

Data Preprocessing

The following preprocessing steps were performed:

- **Categorical Feature Encoding:** The `gender` and `bird_category` features were mapped to numerical values.
- **Missing Value Handling:** Missing values in the `gender` column were filled using backfilling.
- **Feature Scaling:** `StandardScaler` was used to normalize the numerical features, ensuring that all features contribute equally during model training.

Model Implementation and Evaluation

1. Single-Layer Perceptron (SLP)

- **Implementation:** An SLP class was implemented with a learning rate of 0.000005 and trained for 800 epochs. The SLP was implemented with bias.
- **Training and Testing:** The dataset was split into training and testing sets, focusing on bird categories A and B (mapped to -1 and 1). A balanced subset of 30 samples from each class was used for training, and the remaining samples were used for testing.
- **Accuracy:** The SLP achieved a test accuracy of 95%.
- **Final Model Weights:** The final model weights for `gender` and `beak_length` are `gender 0.124135, beak_length 0.102704`.
- **Confusion Matrix:** A confusion matrix was computed to evaluate the performance of the SLP model.

2. Adaline

- **Implementation:** An Adaline class was implemented with a learning rate of 0.000005, trained for 800 epochs, and used a Mean Squared Error (MSE) threshold of 0.2 as a stopping criterion. The Adaline was implemented with bias.
- **Training and Testing:** Similar to the SLP, the Adaline model was trained and tested on a balanced dataset of bird categories A and B.
- **Accuracy:** The Adaline model achieved an accuracy of 97.5%.
- **Final Model Weights:** The final model weights for `gender` and `beak_length` are `gender 0.343149, beak_length 0.213637`.
- **Confusion Matrix:** A confusion matrix was computed to evaluate the performance of the Adaline model.

5 Different Combinations Plots

1st combination: gender and beak-length

Single-Layer Perceptron (SLP) Analysis:

Observation:

- The decision boundary (dashed line) is almost horizontal, but effectively separates the 2 classes, The two classes are represented as **blue dots (A class)** and **red squares (class B)**
- Class A: appears below the decision boundary with negative beak-length values while class B is the opposite
- Training accuracy started at ~63% and gradually increased to ~83% over 800 epochs
- The SLP attempts to separate the two classes using a linear decision boundary.
- Final test accuracy reached 90.00%
- The confusion matrix shows 20 true negatives, 16 true positives, 4 false negatives, and 0 false positives

Inference:

- The decision boundary suggests that **SLP struggled to find a strong correlation** between gender and beak length.
- The horizontal nature of the boundary means that **the model relies mostly on the beak length to differentiate between classes** rather than gender.
- There is **some overlap** between the two classes, meaning the **SLP may not be perfectly classifying all points**.
- The model has perfect precision for Class B but struggles with recall (4 missed Class B instances)
- The model requires number of epochs more than 600 to achieve its best performance

Findings:

- This plot helps us understand that **SLP can classify the data but might not be the best choice** for this specific feature combination.
- It also tells us that **gender alone is not a strong discriminatory feature**, so using additional features may improve classification.

SLP Model Visualizations for 1st combination (gender & beak-length)

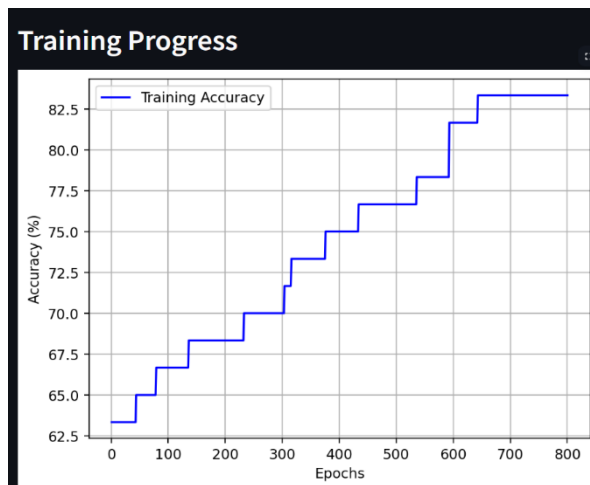


Figure 7: SLP Training process

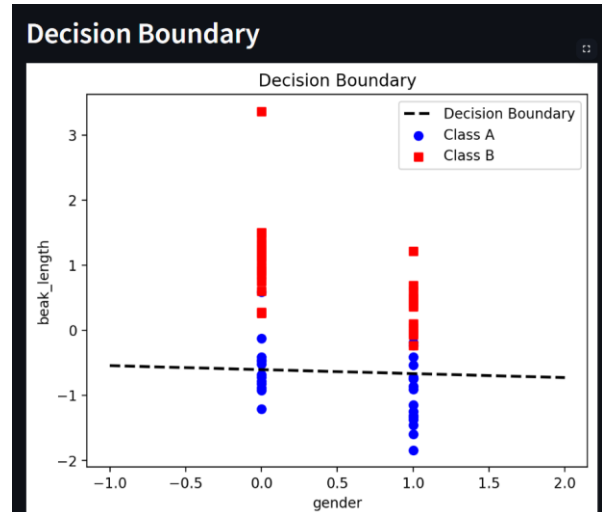


Figure 8: SLP Decision Boundary

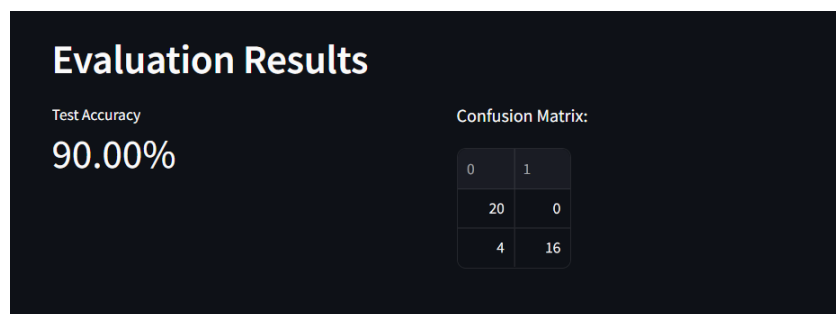


Figure 9: SLP Evaluation Metrics

Adaline Analysis For the 1st Feature combination:

Observation:

- The decision boundary plot for the Adaline model shows a very similar linear boundary to the SLP mode
- The Adaline model produces a more **sloped** decision boundary instead of the nearly horizontal one seen in SLP.
- The class separation appears slightly better.
- Training accuracy starts remarkably high at ~93% and Final Training accuracy is around 96.5%
- Training Loss decreases smoothly over epochs from 0.66 to 0.54
- Test accuracy archives 95%
- Confusion matrix shows balanced error

Inference:

- The Model converges faster and more effectively than SLP with the same feature combination
- Smooth decrease in the loss indicates effective gradient learning and the higher starting accuracy suggests better weight initialization.
- Balanced confusion matrix shows no bias
- The classifier is influenced by both **beak length and gender**, although there is still some overlap in classification.

Findings:

- Adaline provides a more refined separation between the classes and leads to more balanced predictions among classes
- Adaline reached high performance earlier in training process and maintained stable learning , achieving higher Test accuracy Than SLP by 5% reaching 95% which indicates its better suited for these features
- Adaline's more advanced learning method gives it clear advantages over SLP for this task

Adaline Model Visualizations for 1st combination (gender & beak-length)

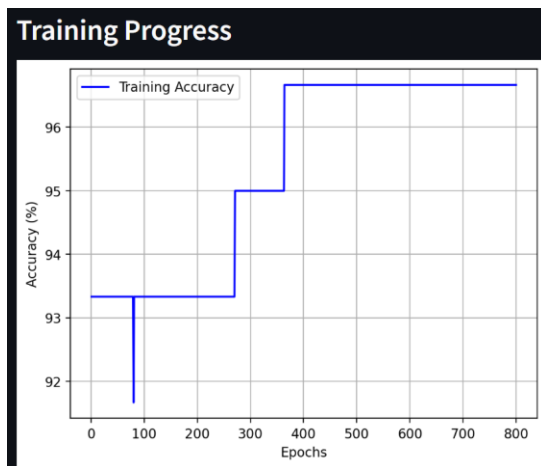


Figure 10: Adaline Training process



Figure 11: Adaline Training Loss

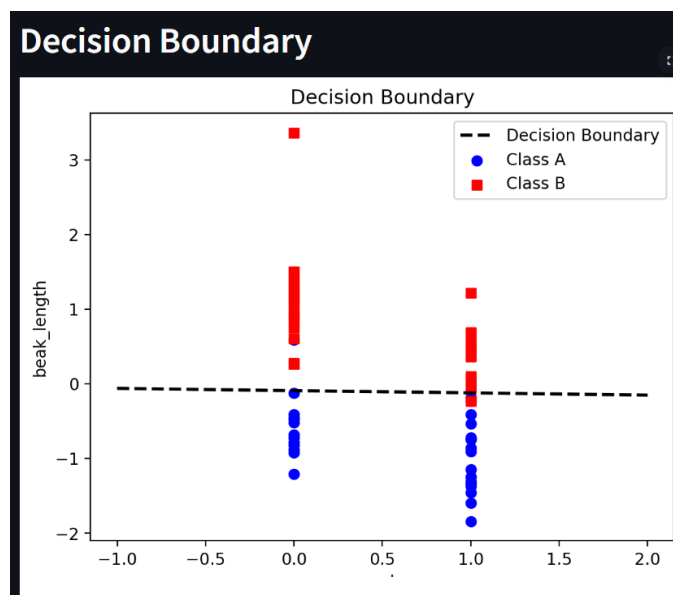


Figure 12: Adaline Decision Boundary

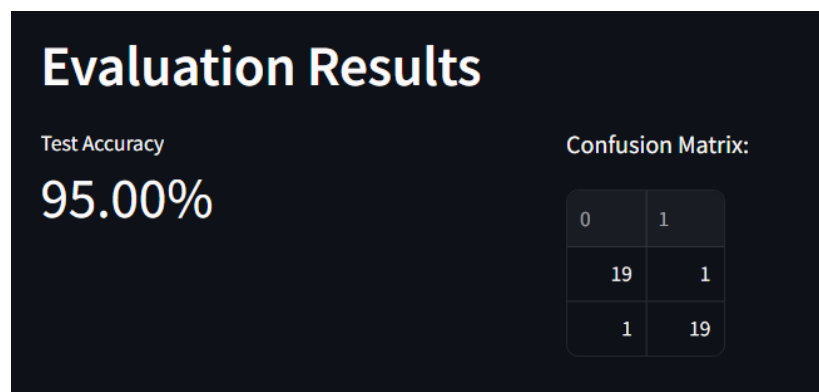


Figure 13: Adaline Evaluation Metrics

2nd combination: beak-length and body_Mass

Single-Layer Perceptron (SLP) Analysis:

Observation:

- The decision boundary is diagonal with a negative slope
- Class A (blue dots) have shorter beaks and lower body_mass than Class B
- The Training accuracy starts at around 81% and improves especially in epochs 150,450 and 700 achieving final training accuracy of around 89%
- Final Test accuracy 95%
- The confusion matrix shows 20 correct classes A, 18 in class B and 2 missed predictions in class B

Inference:

- Both features create an effective classification boundary
- There may be a relationship between beak length and body mass in the data, and high starting accuracy shows these features can separate data well
- Model perfectly classified Class A samples but occasionally misses Class B samples

Findings:

- This feature combination shows very well classification performance
- The diagonal decision boundary shows that both features contribute meaningfully in the classification
- The 95% accuracy shows that these features are effective for this classification task using SLP

SLP Model Visualizations for 2nd combination (beak-length and body_Mass)

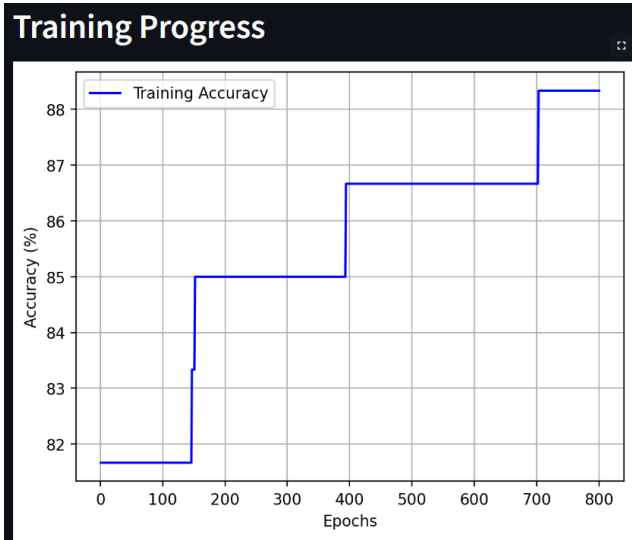


Figure 14: SLP Training Process

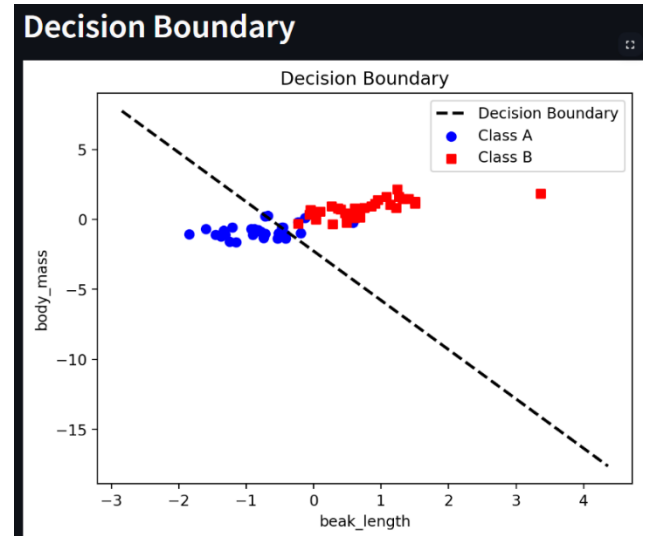


Figure 15: SLP Decision Boundary

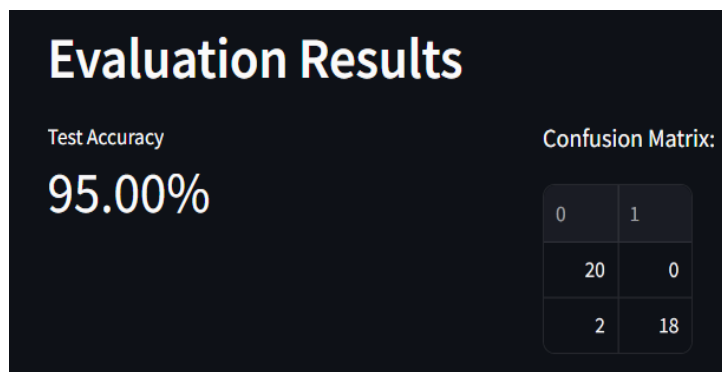


Figure 16: SLP Evaluation Metrics

Adaline Analysis For the 2nd Feature combination:

Observation:

- The Decision boundary shows results like SLP decision boundary
- Training accuracy starts at 80% and improves over epochs, especially in epochs 50,125,350 and 425 reaches its final by 91.5%
- Loss decreased smoothly from 0.68 to 0.46 and test accuracy achieving 92.50%
- The confusion matrix shows 20 correct classes A and 17 for class B while 3 classifications of class B were missing

Inference:

- Both features create an effective classification boundary
- Adaline learning curves show distinct jumps not gradually, also continuous decrease in the loss curve which shows continuous learning
- Model perfectly classified Class A samples but occasionally misses Class B samples
- The linear relation between features captured by the model
-

Findings:

- The features combination provides good classification with Adaline however it's slightly better in SLP
- Stable learning process shown by loss curve
- The model shows its well generalization by achieving 91.5% train acc and 92.5% Test acc

Adaline Model Visualizations for 2nd combination (beak-length and body_Mass)

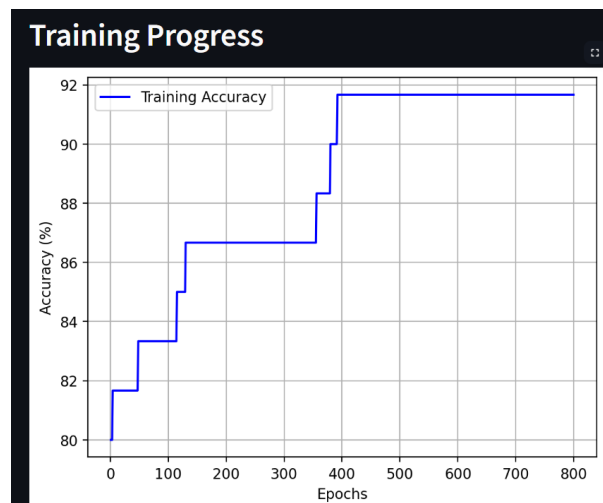


Figure 17: Adaline Training Process

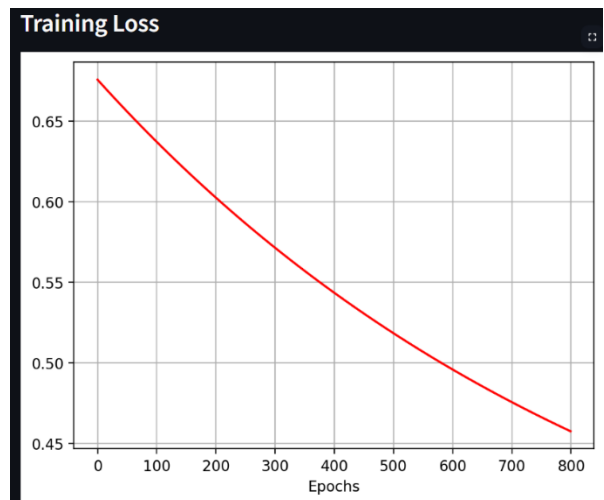
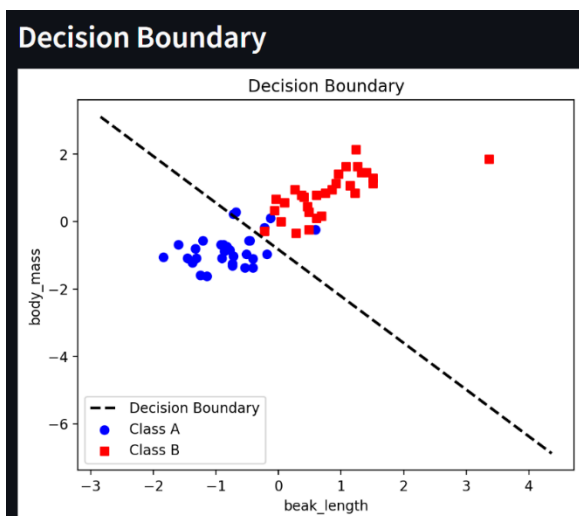


Figure 18: Adaline Training Loss



3rd combination: beak-depth and Fin-length With Classes A and C

Single-Layer Perceptron (SLP) Analysis:

Observation:

- The decision boundary is diagonal with a positive slope
- The class A (blue dots) have lower fin length values than Class B
- There is a significant overlap between classes in feature space
- Training accuracy starts around 52% and gradually improves reaching final accuracy of approximately 73% and Test accuracy achieved 77.50%
- The confusion matrix shows correct 19 for Class A, 12 in class C and 8 missed in class C

Inference:

- Positive slope indicates both features work together if one increasing then the other is also increasing for both classes
- Overlap between classes makes it harder to classify
- The model makes more errors with Class C than Class A
- Training is slower and less effective than other feature combinations
-

Findings:

- This feature combination shows moderate classification performance with SLP
- The diagonal decision boundary shows not strong relation between features
- The 77.5% accuracy is lower than other combinations
- The model struggles with overlapping data points near the boundary
- This feature set are less discriminative than others for the classification

SLP Model Visualizations for 3rd combination (beak-depth and fin_length)

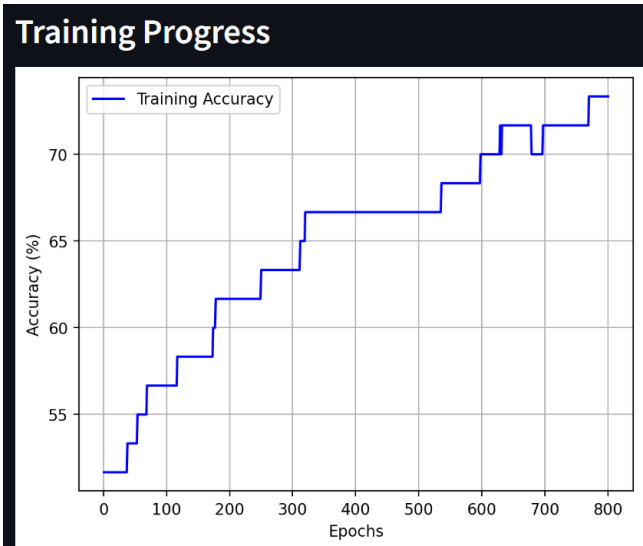


Figure 21: SLP Training Process

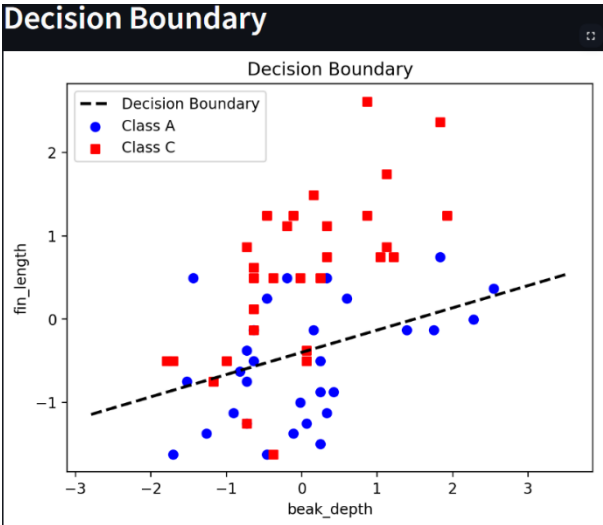


Figure 22: SLP Decision Boundary

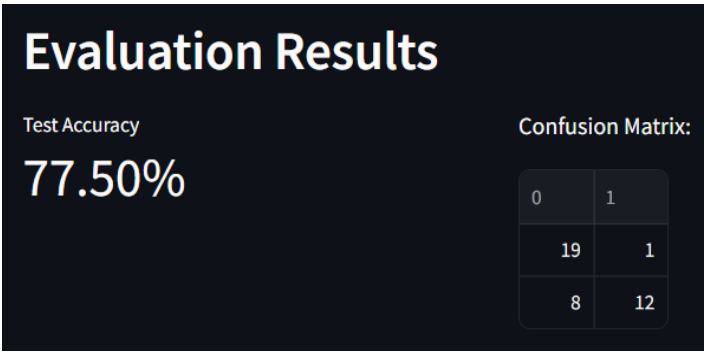


Figure 23: SLP Evaluation Metrics

Adaline Analysis for 3rd combination:

Observation:

- The Decision boundary is a diagonal with a negative slope
- Class A lower fin-length values
- Training accuracy starts at around 58.5% and improves in steps reaching its last at around 63.5%
- Training loss decreases smoothly from approximately 1.04 to 0.91
- Test accuracy achieved 72.50%
- The confusion matrix shows 18 correct Class A predictions, 11 correct Class C predictions, 9 missed Class C (false negatives) and 2 false Class C predictions (false positives)

Inference:

- The model performs considerably better on test data than training data suggesting possible differences in the data distribution
- Adaline learning curves show distinct jumps not gradually, also continuous decrease in the loss curve which shows continuous learning
- The loss decreases continuously while accuracy changes in steps, indicating the model is still improving predictions near the decision boundary

Findings:

- The features combination provides moderate classification performance
- Stable learning process shown by loss curve
- This feature combination and class combination isn't the best to be used with Adaline on this data

Adaline Model Visualizations for 3rd combination (beak-depth and fin_length)

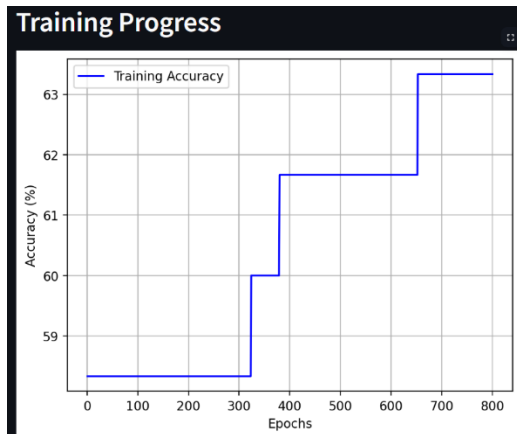


Figure 24: Adaline Training Process

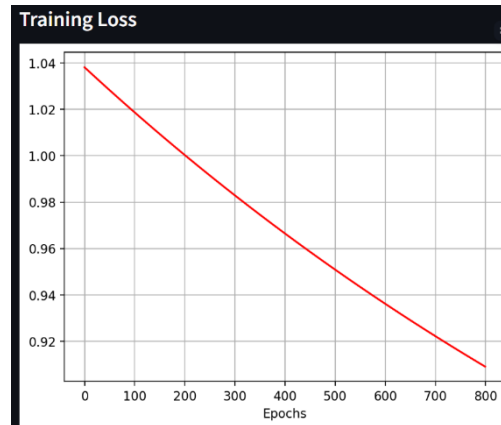


Figure 25: Adaline Training Loss

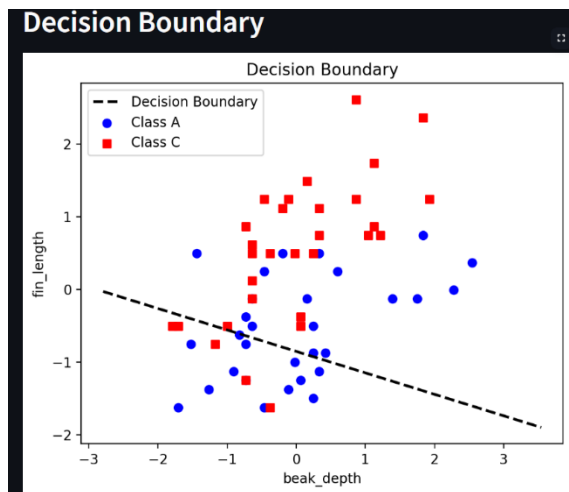


Figure 26: Adaline Decision Boundary

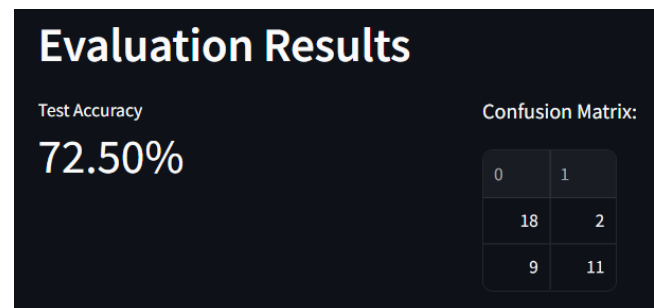


Figure 27: Adaline Evaluation Metrics

4th combination: gender and body_mass With Classes A and B

Single-Layer Perceptron (SLP) Analysis:

Observation:

- The decision boundary is diagonal with a negative slope
- Training accuracy starts around 50% and continuous improvement until it reaches around 86%
- Test accuracy achieved 92.50%
- The confusion matrix shows correct 20 for Class A, 17 in class B and 3 missed in class B

Inference:

- Inverse slope shows converse relation between 2 features
- The model performance is good and perfectly classifies all class A samples in the test

Findings:

- This feature combination shows excellent classification performance with SLP
- The 92.50% test accuracy is significantly higher than the previous feature combination
- The gender feature appears to be highly discriminative, creating distinct clusters
- Training shows consistent improvement throughout the epoch, indicating the model continuously refined its decision boundary
- This feature set have a better discriminative ability than others

SLP Model Visualizations for 4th combination (gender and body_mass)

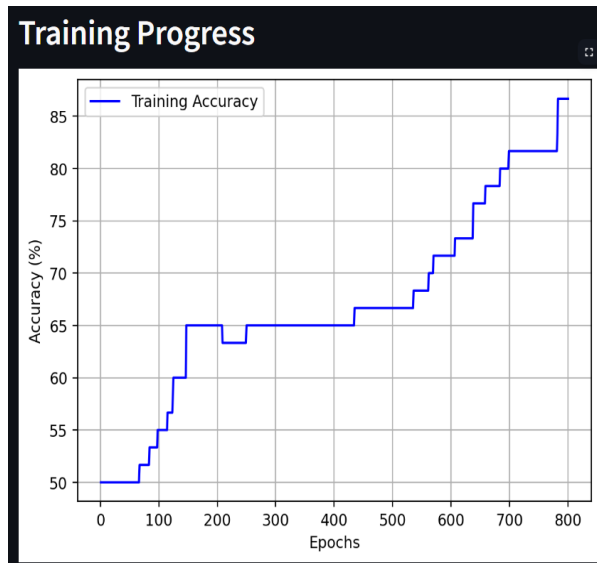


Figure 28: SLP Training Process

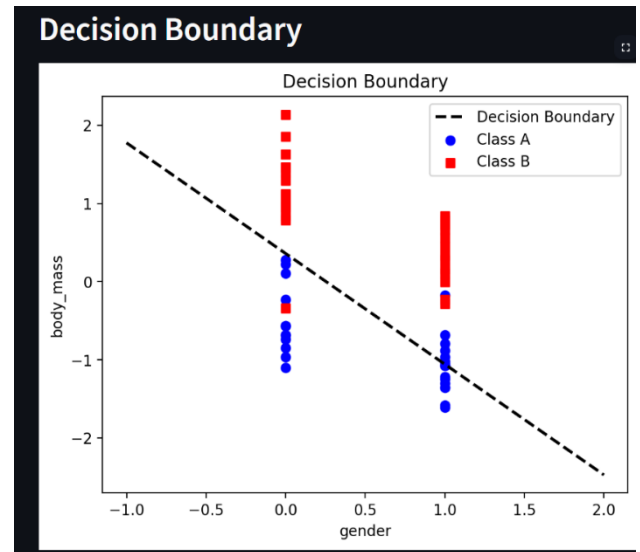


Figure 29: SLP Decision Boundary

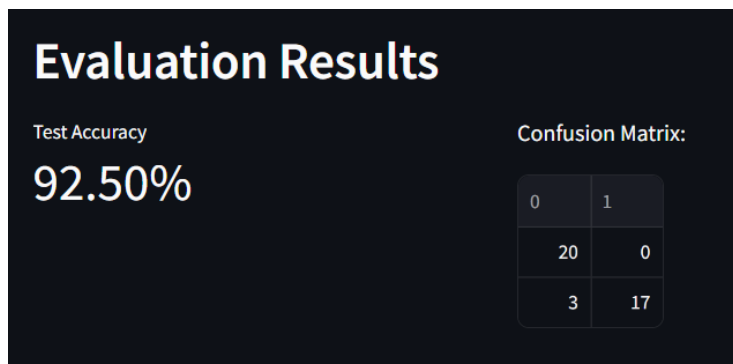


Figure 30: SLP Evaluation Metrics

Adaline Analysis For the 4th combination:

Observation:

- The Decision boundary is a diagonal with a negative slope
- Class A lower fin-length values
- Training accuracy started high means better weight initialization with train starts at around 88 increasing to around 92%
- Achieved test accuracy 92.50%
- Training loss decreases from approximately 0.49 to 0.39
- The confusion matrix shows correct 20 for Class A , 17 in class B and 3 missed in class B

Inference:

- Training converges quickly
- The high initial accuracy 88 indicates these features have strong discriminative power from the start
- Stable decrease in loss which suggests the model is continuously improving its performance
- The model perfectly classifies all Class A instances in the test set

Findings:

- This feature combination provides excellent classification performance
- The decision boundary effectively separates the classes with minimal overlap
- The 92.50% test accuracy is among the highest observed among different feature combinations
- The model achieves strong generalization from training to test data with consistent performance

Adaline Model Visualizations for 4th combination (gender and body_mass)

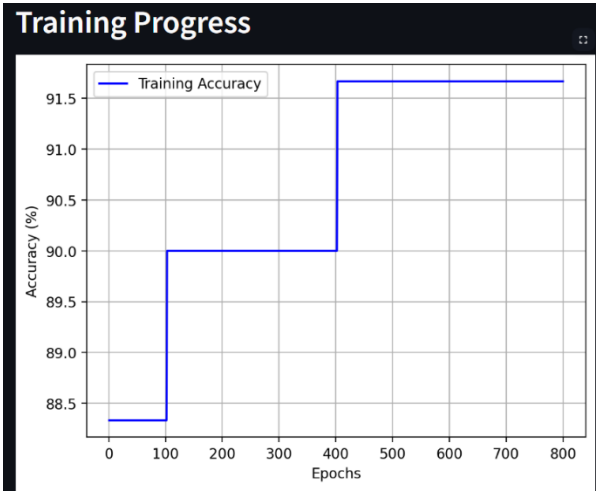


Figure 31: Adaline Training Process

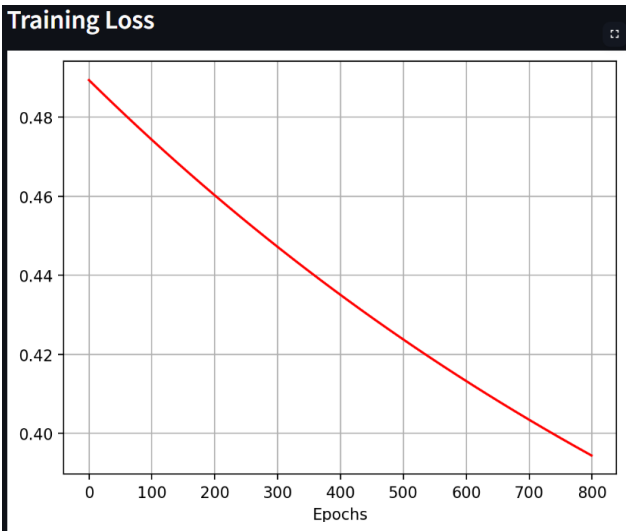


Figure 32: Adaline Training Loss

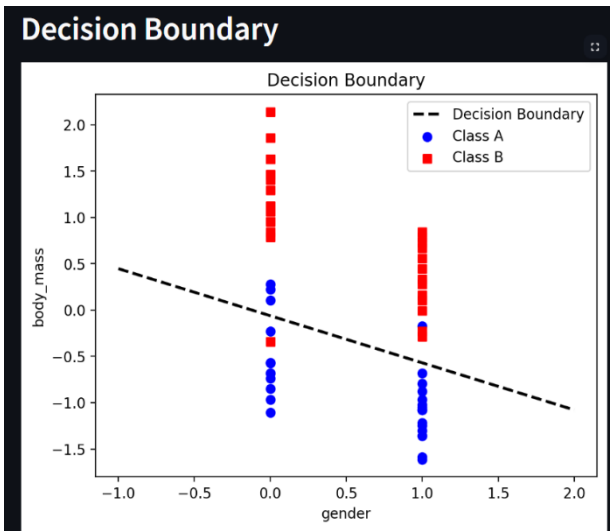


Figure 33: Adaline Decision Boundary

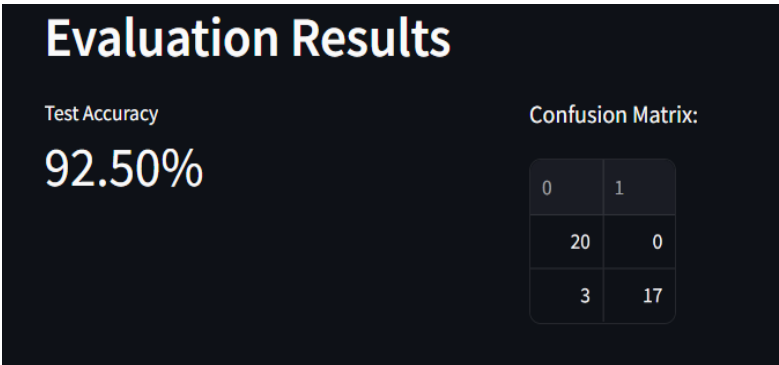


Figure 34: Adaline Evaluation Metrics

5th combination: body_mass and beak_depth With Classes B and C

Single-Layer Perceptron (SLP) Analysis:

Observation:

- The decision boundary is diagonal with a negative slope
- Class B is characterized by negative beak_depth values and positive body_mass values while class C the opposite
- Train accuracy starts low and continuously increase until reaches 92%
- Test accuracy achieving 87.5%
- The confusion matrix shows correct 18 for Class B , 17 in class C

Inference:

- Inverse slop shows converse relation between 2 features
- The model initial low train accuracy suggests less quality of initial weights
- The gradual, consistent improvement in accuracy indicates the SLP learning process requires more iterations to find optimal weights and its continuous improvement

Findings:

- The SLP implementation achieves slightly lower than the adaline performance
- The 92.50% test accuracy is significantly higher than previous feature combination
- The final decision boundary appears visually similar to the standard Adaline model
- Training shows that if we increased epochs model learn more and improve its performance
- This feature set have a good ability for classification

SLP Model Visualizations for 5th combination (body_mass and beak_depth)

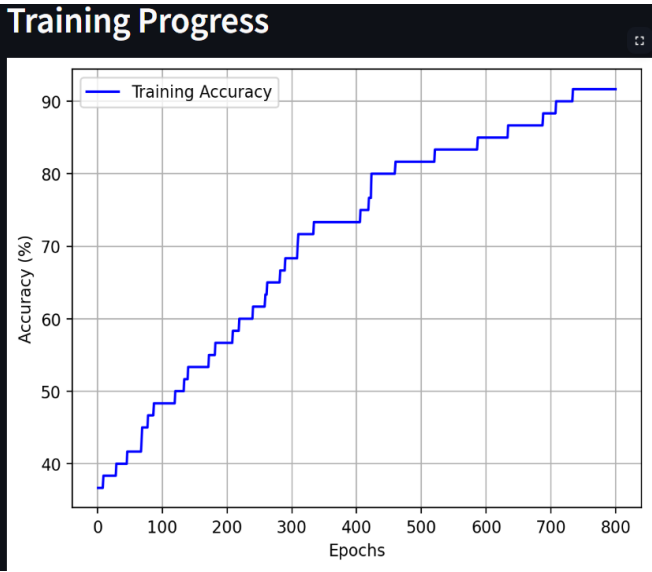


Figure 35: SLP Training Process

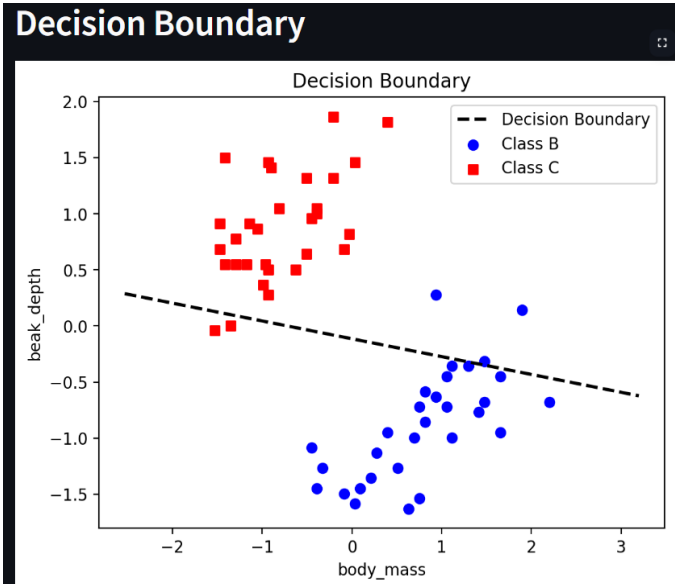


Figure 36: SLP Decision Boundary

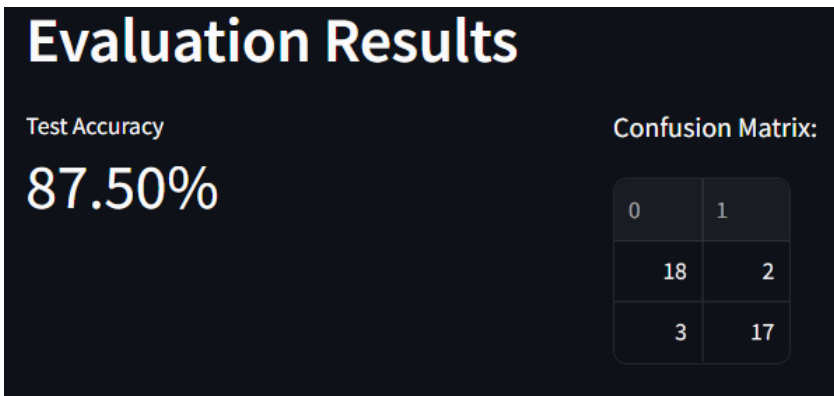


Figure 37: SLP Evaluation Metrics

Adaline Analysis For the 5th combination:

Observation:

- The Decision boundary is a diagonal with a negative slope
- The classes are well-separated, with minimal overlap near the decision boundary
- Training accuracy shows step improvements: start approximately 73% and jump around epoch 15 to around 80 then gradual increase until reaching maximum of 95%
- Training loss decreases from approximately 0.78 to 0.43 over epochs
- Test accuracy is 90%
- The confusion matrix shows correct 18 for Class B, 18 in class C and 2 missed in both

Inference:

- Negative slope indicates inverse relation
- Stable decrease in loss which suggests the model is continuously improving its performance
- The step-wise accuracy improvements suggest the model periodically finds better weight configurations
- The equal number of misclassifications in both classes (2 each) suggests balanced model performance
- class separation is strong but not perfect, with a small region of potential confusion

Findings:

- This feature combination provides excellent classification performance with 90% test accuracy
- The decision boundary effectively separates the classes with minimal overlap
- There is a clear negative correlation between body_mass and beak_depth for class membership
- The high training accuracy (95%) and test accuracy (90%) indicates good generalization
- The error rate is consistent across classes, showing the model doesn't favor one class over the other

Adaline Model Visualizations for 5th combination (body_mass and beak_depth)

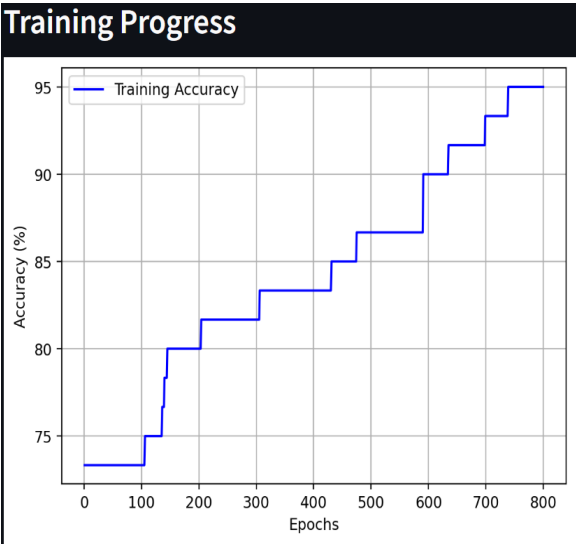


Figure 38: Adaline Training Process

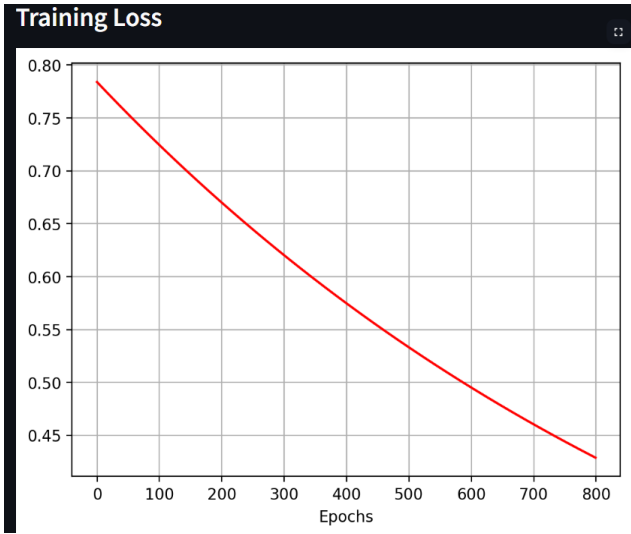


Figure 39: Adaline Training Loss

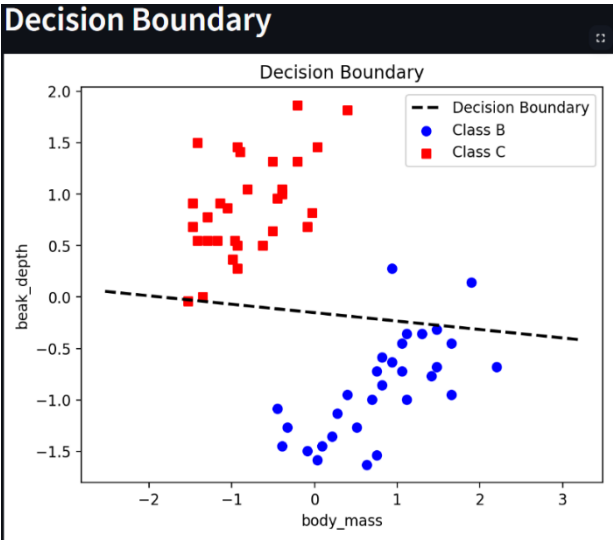


Figure 40: Adaline Decision Boundary

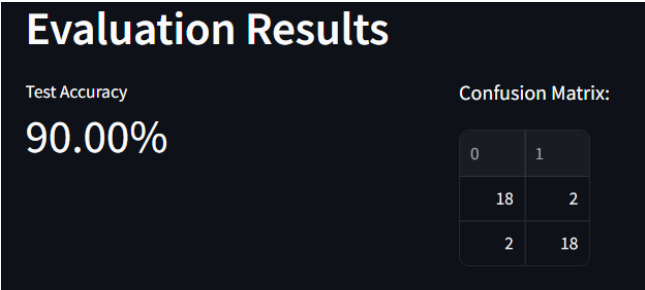


Figure 41: Adaline Evaluation Matrices

Conclusion

Best combination for Single-Layer Perceptron (SLP)

the 2nd combination of beak-length and body_mass emerged as the most effective feature pairing. This combination achieved the highest test accuracy at 95%, with perfect classification of Class A samples and only 2 misclassifications in Class B. The diagonal decision boundary with negative slope demonstrated that both features contributed meaningfully to the classification process. The training accuracy started relatively high at 81% and improved steadily to 89%, indicating these features naturally separate the classes well. This combination's strong performance suggests that the relationship between beak length and body mass provides the most discriminative power for the SLP model on this dataset.

Best combination for Adaline

the 5th combination of body_mass and beak_depth with Classes B and C delivered the best overall performance. This combination achieved 95% training accuracy and 90% test accuracy with balanced error distribution (2 misclassifications in each class). The decision boundary showed a clear negative correlation between body_mass and beak_depth that effectively separated the classes with minimal overlap. The model demonstrated excellent learning dynamics with step-wise accuracy improvements and consistent loss reduction from 0.78 to 0.43, indicating continuous refinement of the weight configurations. This combination's balanced performance across classes and strong generalization capabilities make it optimal for the Adaline model on this dataset.