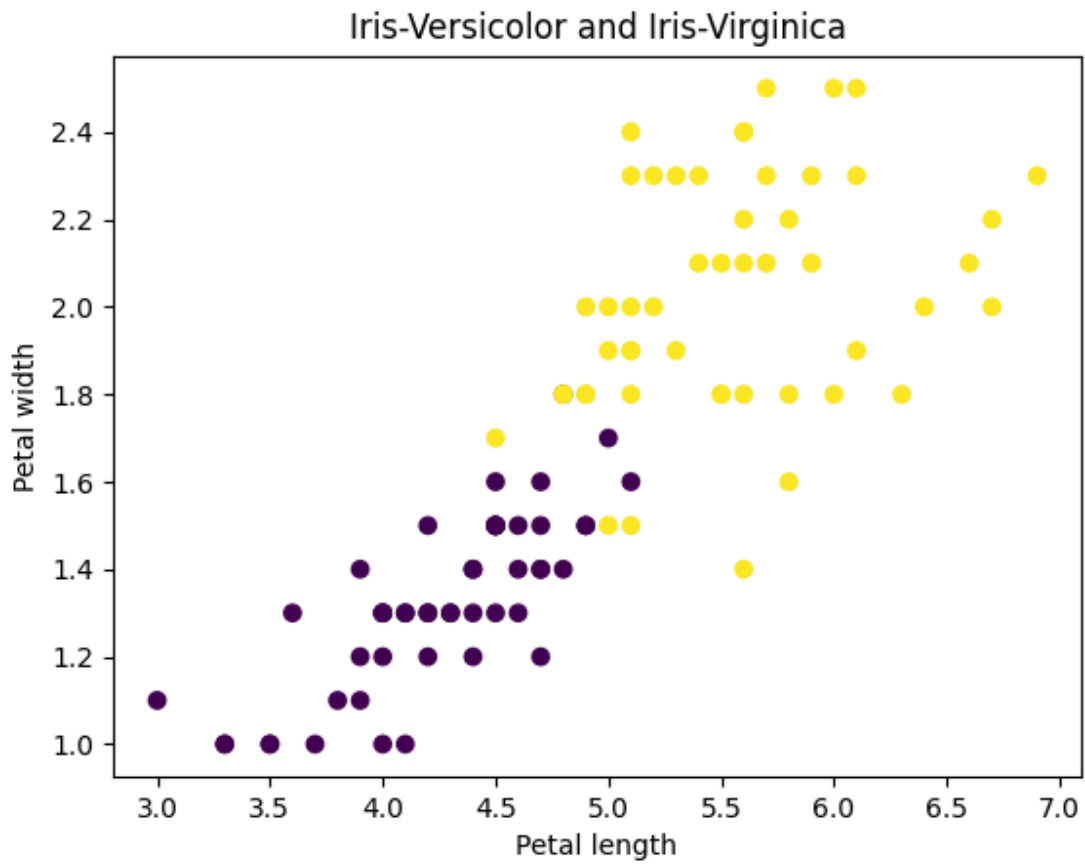
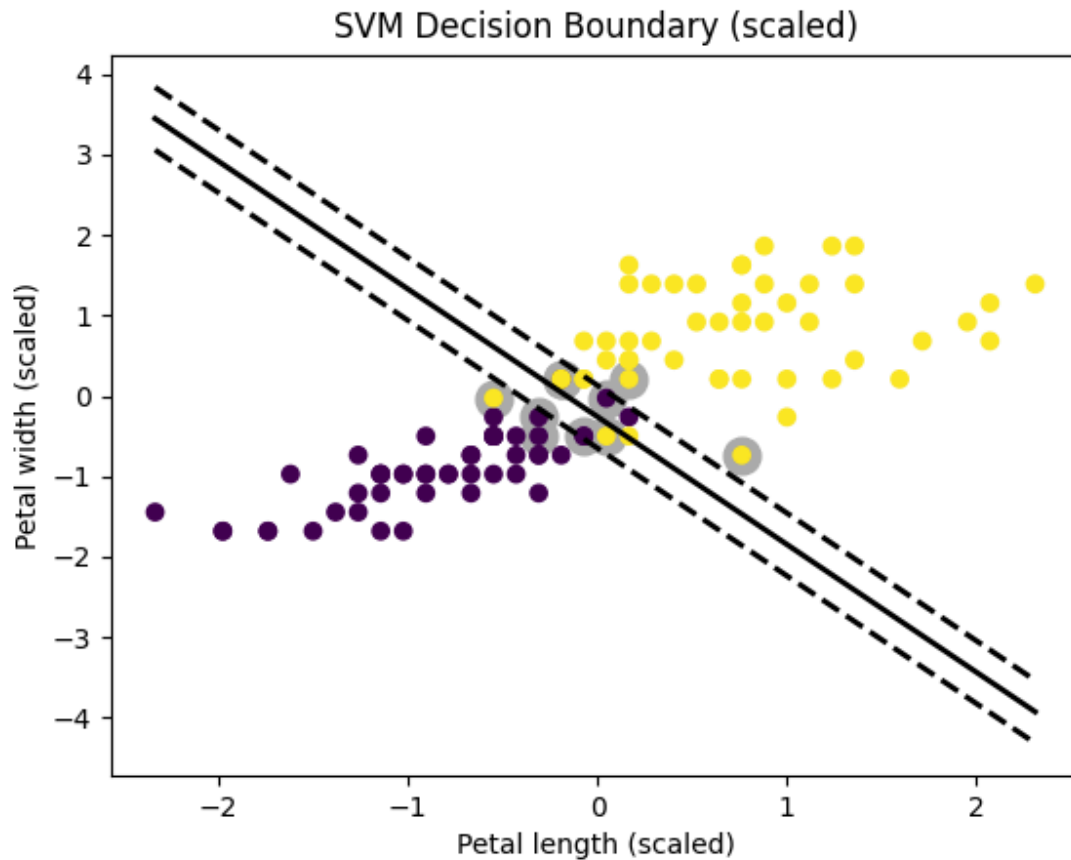


DAMA61 - Final Exam - Exercise 1

Output:



DAMA61 - Final Exam - Exercise 1



SVM Model with $C=10$: Train accuracy=0.95, Test accuracy=0.95

SVM Model with $C=50$: Train accuracy=0.95, Test accuracy=0.9

SVM Model with $C=100$: Train accuracy=0.95, Test accuracy=0.9

SVM: Train accuracy=0.95, Test accuracy=0.95

Logistic Regression: Train accuracy=0.95, Test accuracy=0.9

Decision Tree: Train accuracy=0.975, Test accuracy=0.95

AdaBoost: Train accuracy=0.95, Test accuracy=0.8

Probability of being Iris Versicolor for SVM: 0.4803

Probability of being Iris Versicolor for Logistic Regression: 1.0000

Probability of being Iris Versicolor for Decision Tree: 1.0000

Probability of being Iris Versicolor for AdaBoost: 0.9969

DAMA61 - Final Exam - Exercise 1

Comments:

The SVM model with $C=10$ achieved a very good train and test accuracy of 0.95. Increasing the regularization parameter to 50 and 100 resulted in slightly lower test accuracy (0.9), suggesting that these values may lead to overfitting. The decision boundary plot for the SVM model with $C=10$ shows a clear separation between the two classes with correctly identified support vectors.

Logistic Regression performed well with a train accuracy of 0.95 and a test accuracy of 0.9. This method is quite effective for this binary classification task.

The probability of the sample being Iris Versicolor is 1.0000, indicating that the Logistic Regression model has high confidence in this prediction.

The Decision Tree model with a maximum depth of 2 achieved a train accuracy of 0.975 and a test accuracy of 0.95. This shows that even a simple Decision Tree can perform well on this task.

The probability of the sample being Iris Versicolor is 1.0000, which indicates a high level of confidence.

The AdaBoost model had a lower test accuracy of 0.8, indicating it might not be as suitable for this specific dataset or the number of estimators could be insufficient.

The probability of the sample being Iris Versicolor is 0.9969, which is very high, showing a strong prediction, but it has the lowest test accuracy among the classifiers, suggesting potential overfitting.