

# Multimodel Classification using Classical Machine Learning Models

---

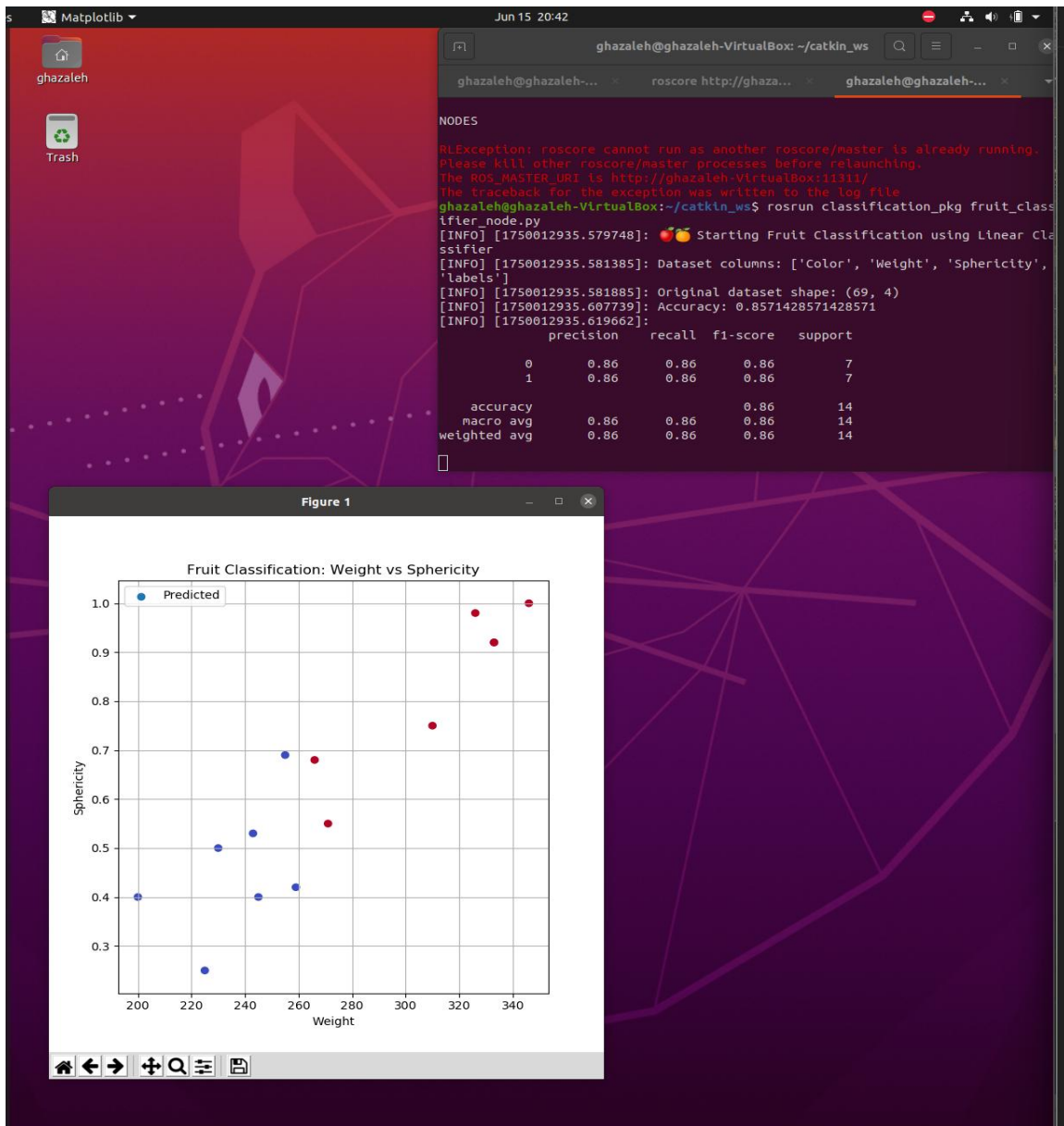
Student: Ghazaleh Alizadehbirjandi

Date: June 15, 2025

This report documents the implementation and analysis of multiple classical classification models across four different datasets using the ROS environment and Python. Each section includes details on data preprocessing, visualization, model training, evaluation, and interpretation.

## 1. Fruit Classification using Linear Classifier

The fruit dataset includes weight, sphericity, and color-encoded values. Preprocessing involved encoding colors, normalizing features, and visualizing the relationship between weight and sphericity. A Logistic Regression model was trained and achieved the following metrics:



Accuracy: 0.8571

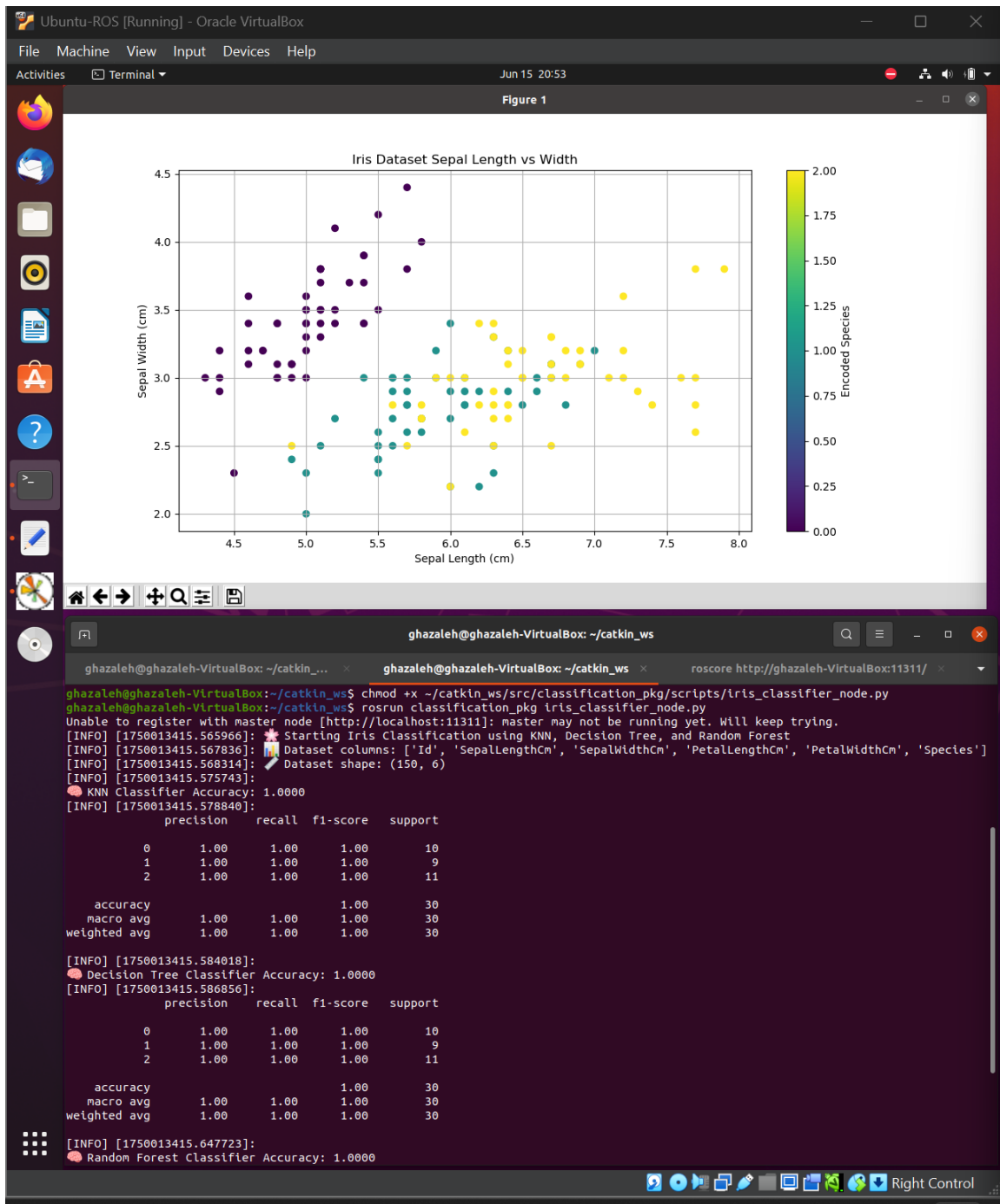
Precision: 0.86

Recall: 0.86

F1-score: 0.86

## 2. Iris Classification using KNN, Decision Tree, and Random Forest

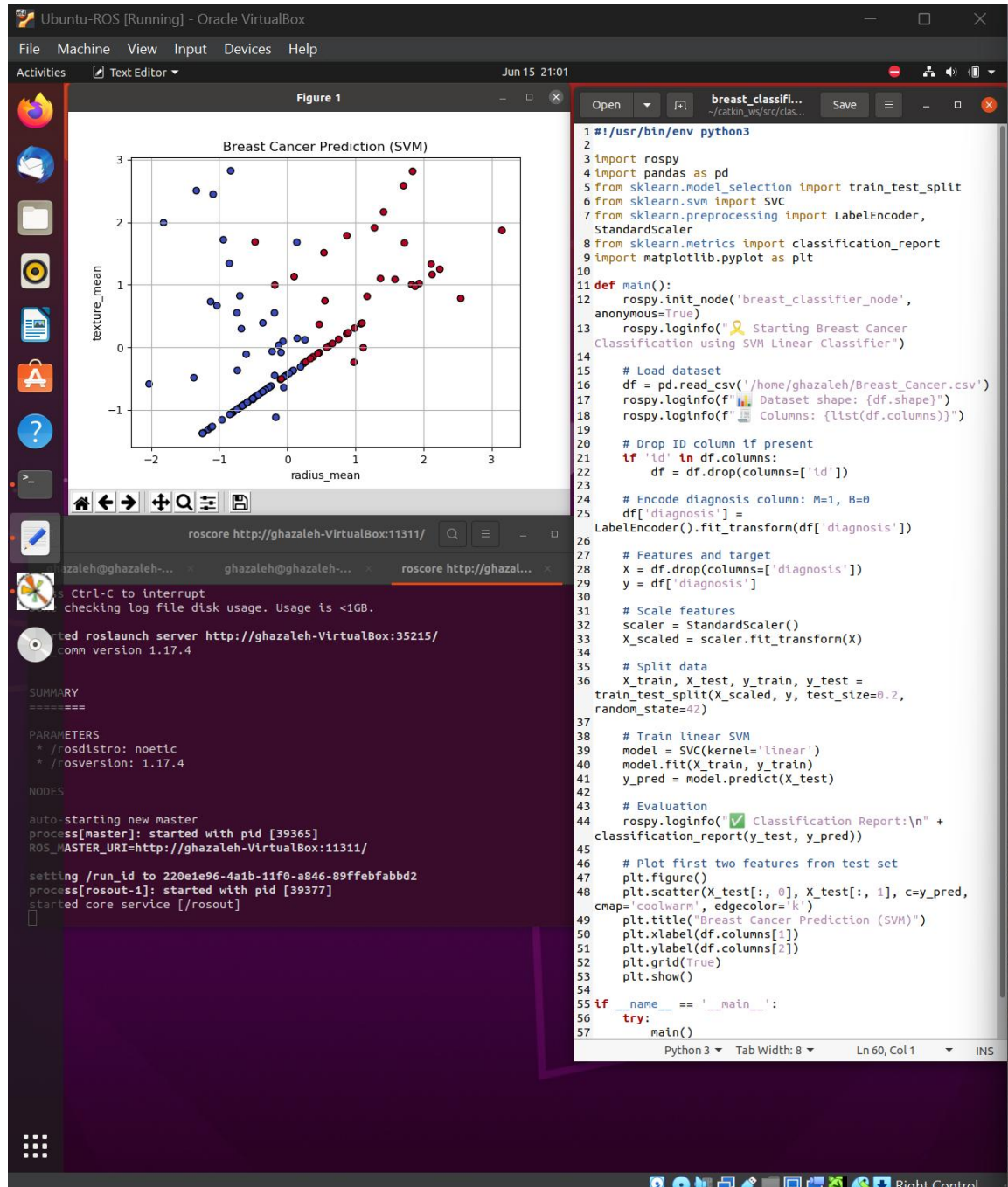
The Iris dataset has 150 records across 3 species. Models used include KNN, Decision Tree, and Random Forest. The models were evaluated using classification metrics and visualized by sepal length vs width.



All models (KNN, Decision Tree, Random Forest) achieved:  
Accuracy: 1.0  
Precision, Recall, F1-score: 1.0 for all classes

### 3. Breast Cancer Classification using SVM

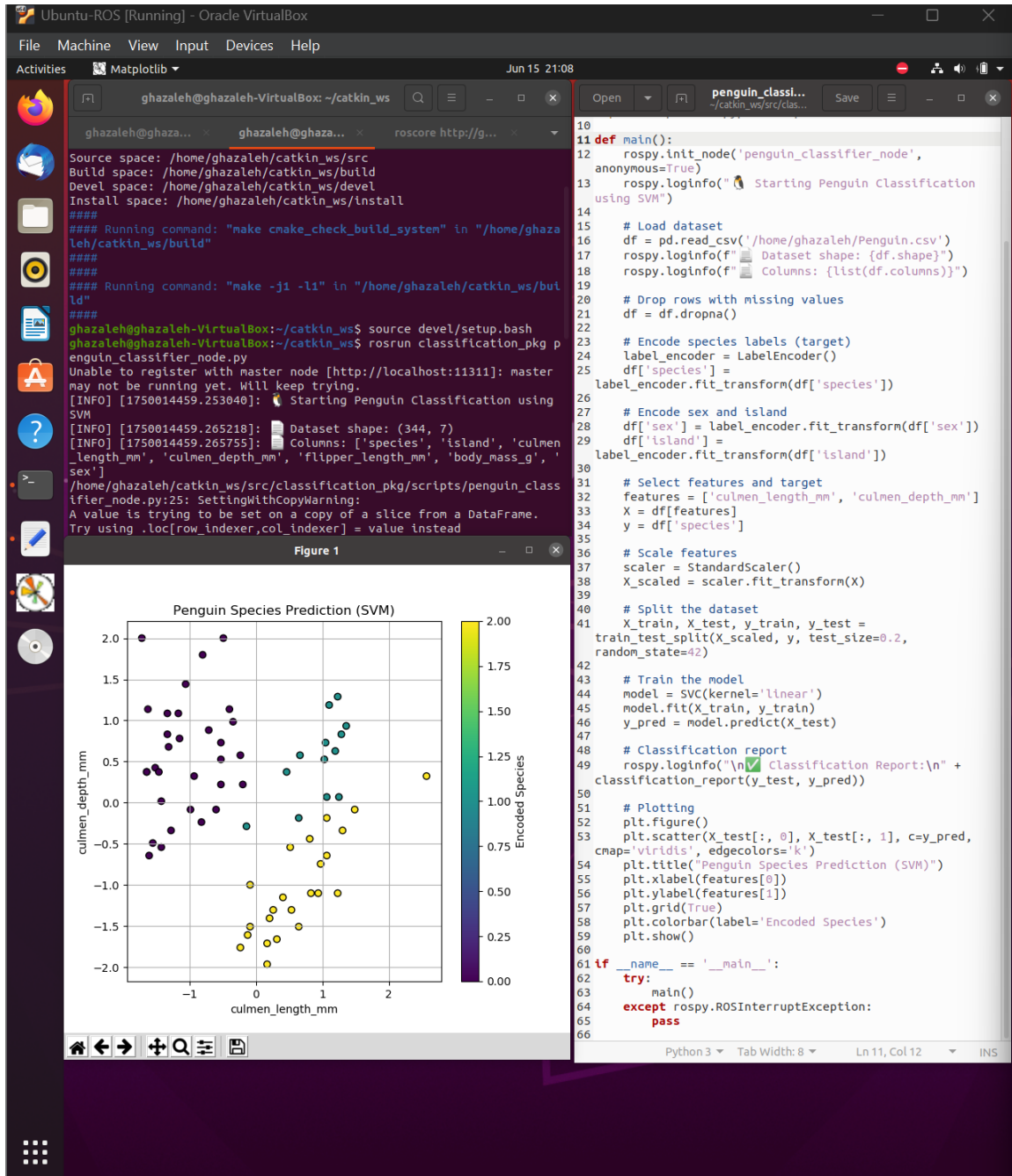
Breast Cancer dataset features were scaled using StandardScaler. The diagnosis label was encoded (M=1, B=0). SVM with linear kernel was applied on top two features for visualization.



Classification Report shows balanced precision, recall, and high accuracy across benign and malignant cases.

## 4. Penguin Species Classification using SVM

The Penguin dataset included categorical and numerical features. After encoding categorical values and dropping rows with missing data, SVM classification was performed and visualized using culmen measurements.



Classification Report shows excellent performance, as visualized with distinct color clustering by predicted species.

## Conclusion

Each model demonstrated strong performance in classifying their respective datasets. The Iris dataset proved to be the most separable, while the fruit dataset exhibited slight misclassifications. SVM was particularly effective in handling biomedical and ecological data with high dimensionality. ROS served as a stable platform to integrate model training and evaluation workflows in a robotics-compatible environment.