📄 Project Report – Iris Flower Classification

1. Introduction

The Iris dataset is one of the most famous datasets in machine learning, commonly used for classification problems. The dataset contains 150 samples of iris flowers with four features: sepal length, sepal width, petal length, and petal width. The goal of this project was to classify the flowers into one of the three species: **Setosa, Versicolor, or Virginica**.

---

2. Methodology

- **Dataset**: Iris dataset (150 samples, 3 classes, 4 numerical features).
- **Preprocessing**: Data was clean with no missing values. Features were used directly for training.
- **Algorithms Used**:
    - Logistic Regression
    - Decision Tree Classifier
    - Random Forest Classifier
- **Evaluation**: Models were trained on 70% of the data and tested on 30%. Performance was measured using accuracy and classification report (precision, recall, F1-score).

---

3. Results

Dataset Info

- Total samples: 150
- Features: Sepal Length, Sepal Width, Petal Length, Petal Width
- Classes: Setosa (50), Versicolor (50), Virginica (50)

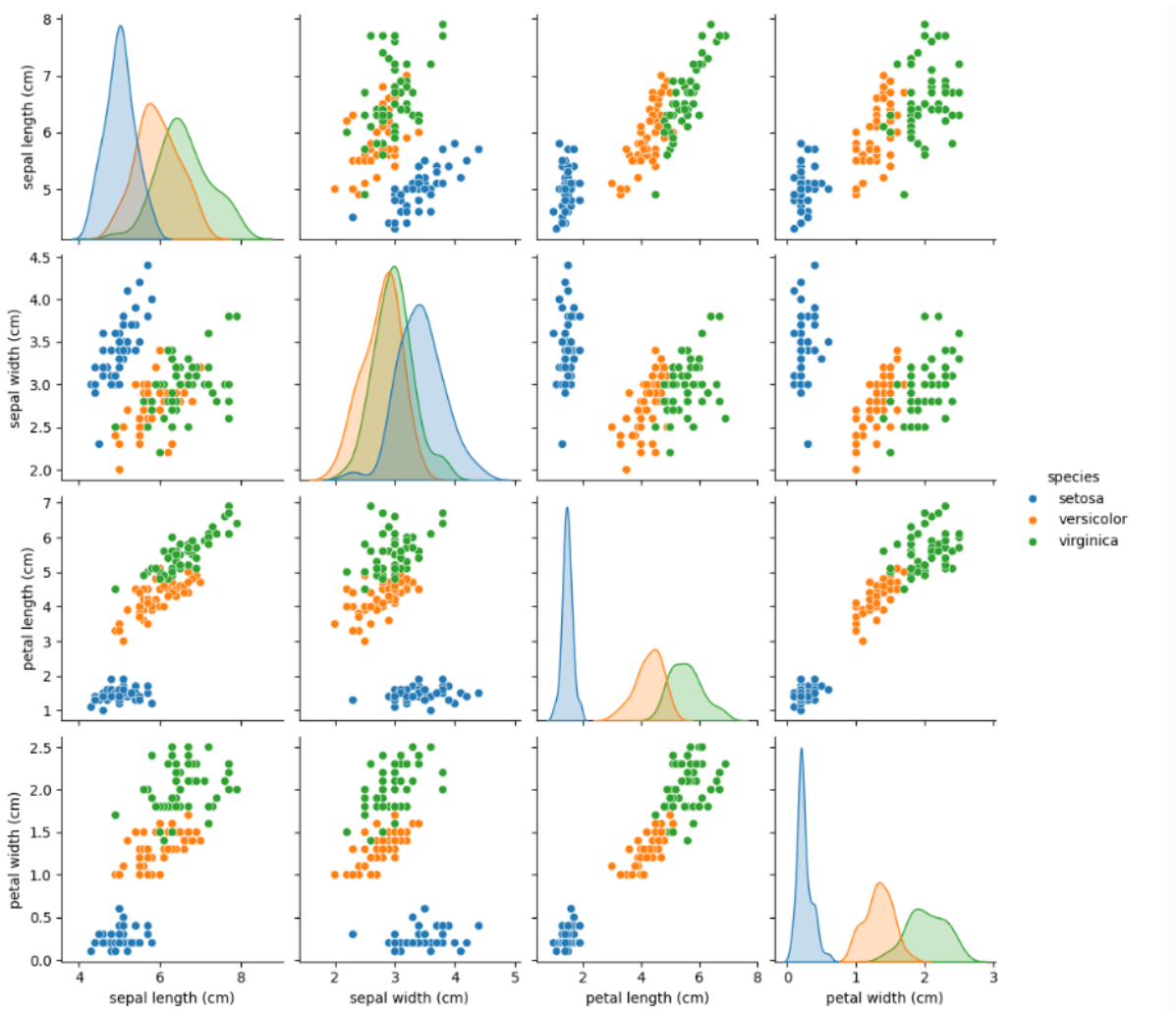
Accuracy Scores

- **Logistic Regression → 100%**
- **Decision Tree → 100%**
- **Random Forest → 100%**

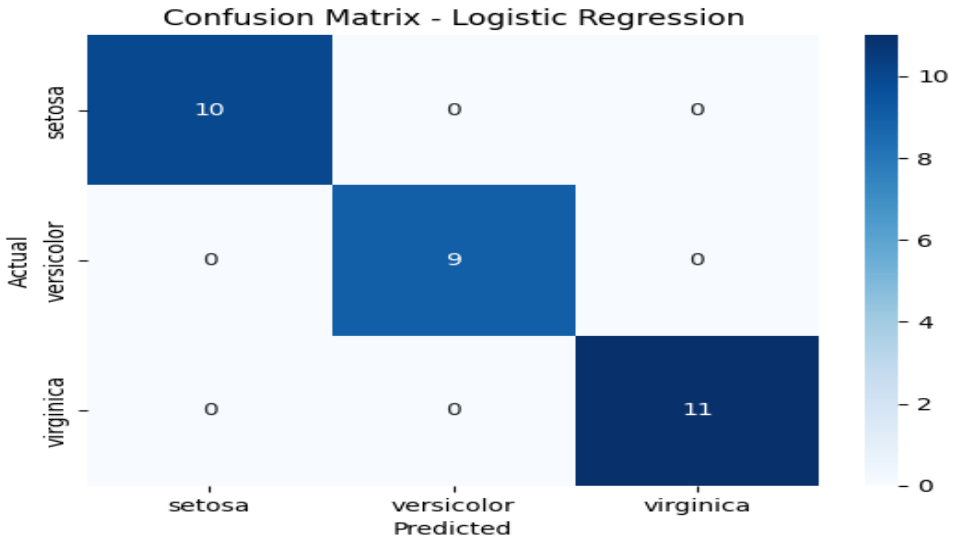
All three models achieved **perfect accuracy** on the test dataset.

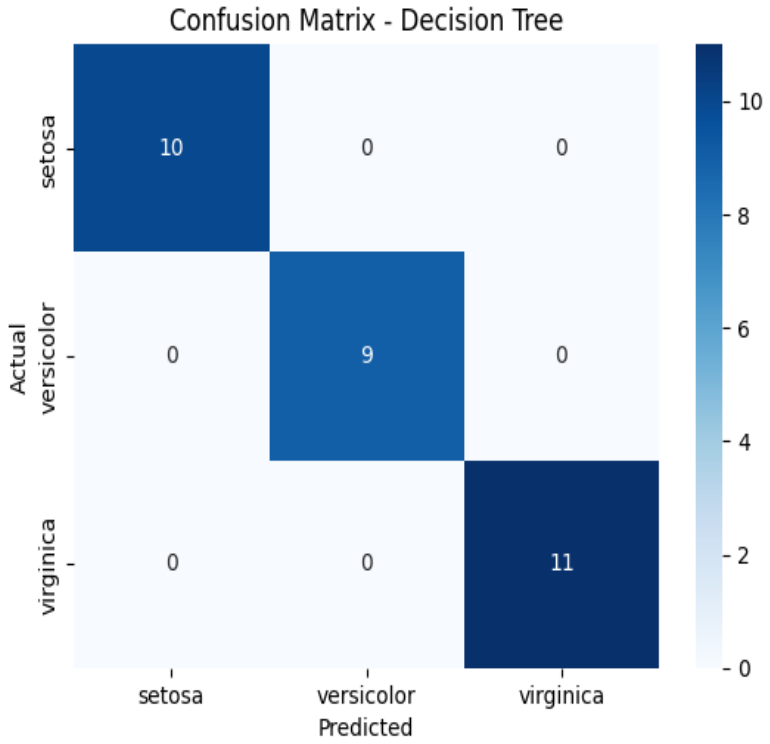
Graphs:

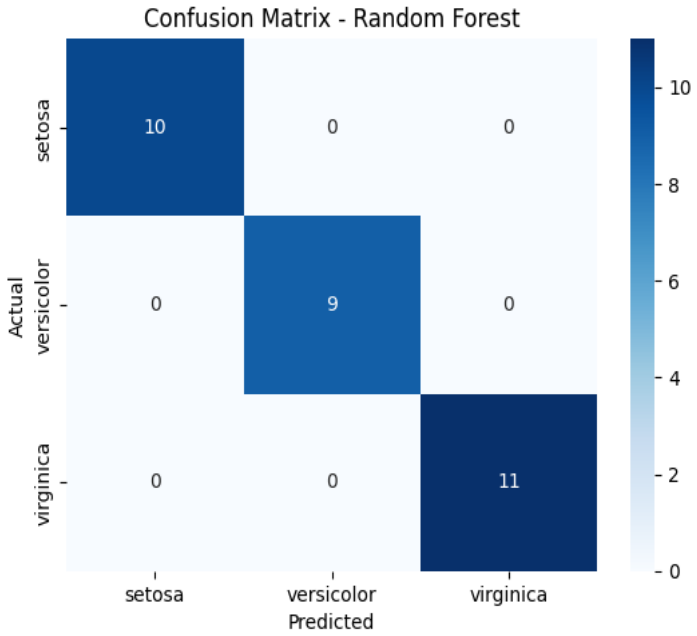
- Pairplot of features (shows class separation).



- Confusion Matrix (all predictions correct).


Confusion Matrix - Logistic Regression
setosa
versicolor
virginica
Actual
10
0
0
0
9
0
0
0
11
setosa
versicolor
virginica
Predicted
10
8
6
4
2
0


Confusion Matrix - Decision Tree
setosa
versicolor
virginica
Actual
10
0
0
0
9
0
0
0
11
setosa
versicolor
virginica
Predicted
10
8
6
4
2
0



---

4. Conclusion

The classification models performed extremely well on the Iris dataset, achieving **100% accuracy** across Logistic Regression, Decision Tree, and Random Forest.

Among them, **Logistic Regression** is selected as the **best model** since it provides perfect accuracy with a simple and interpretable approach.

This project demonstrates the effectiveness of machine learning algorithms in handling multi-class classification problems.