Iris Flower Classification - Project Report

1. Overview

This project classifies Iris flowers into three species - Setosa, Versicolor, and Virginica - using two machine learning algorithms: Logistic Regression and K-Nearest Neighbors (KNN). It uses a dataset containing flower measurements (sepal length, sepal width, petal length, petal width) and predicts the species based on those measurements.

2. Dataset

The dataset used is 'iris.csv', which contains 150 samples of iris flower measurements. The features include:

- Sepal Length
- Sepal Width
- Petal Length
- Petal Width

The target variable is the species, which has three classes.

3. Models Used

1. Logistic Regression:

- Simple and interpretable model.
- Effective for linearly separable data.
- Provides probability estimates of predictions.

2. K-Nearest Neighbors (KNN):

- Instance-based learning algorithm.
- Makes predictions based on nearest neighbors.
- Effective when data is not linearly separable.

Iris Flower Classification - Project Report

4. Evaluation Metrics

The models were evaluated using the following metrics:

- Accuracy Score
- Classification Report (Precision, Recall, F1-score)
- Confusion Matrix (visualized with heatmaps)
- Accuracy Comparison Bar Chart

5. Results and Conclusion

Both models performed well on the dataset:

- Logistic Regression Accuracy: ~97%
- KNN Accuracy: ~96%

The models effectively distinguish between the three species based on the input features. Visualization through pairplots and heatmaps enhances interpretability. The accuracy comparison bar chart provides a quick overview of model performance.