**Iris Species Classification using Logistic Regression**

**1. Introduction**

The **Iris dataset** is a well-known dataset used for classification problems in machine learning. This project aims to classify **three species of iris flowers** (*Setosa, Versicolor, and Virginica*) based on four features:

* **Sepal length**
* **Sepal width**
* **Petal length**
* **Petal width**

We use **Logistic Regression**, a widely used classification algorithm, to build and evaluate a model for predicting the species of iris flowers.

**2. Dataset and Preprocessing**

We loaded the **Iris dataset** from sklearn.datasets, which contains **150 samples** with **four numerical features**. The dataset was converted into a Pandas DataFrame for better handling.

We split the dataset into **80% training data** and **20% testing data** using train\_test\_split(), ensuring class distribution is maintained with **stratified sampling**.

**3. Model Training**

We implemented **Logistic Regression** using LogisticRegression(max\_iter=200) and trained it on the **training dataset**. The model learns to classify iris species based on the provided feature values.

**4. Model Evaluation**

After training, the model was tested on the unseen **test dataset**. The predictions were compared against the actual species labels using **accuracy score**.

* **Predicted Labels:** The model outputs class labels (0, 1, or 2) representing different iris species.
* **Accuracy:** The model achieved an accuracy of **96%**, indicating strong performance in classifying iris flowers correctly.