**Assignment 6 – Supervised Learning: Classification**

Week: 6

Student Name: Bisma Khidmat

Assignment Title: Logistic Regression vs Random Forest Classifier

Tool Used: Google Colab

**Objective**

The goal of this assignment is to apply two supervised learning algorithms — Logistic Regression and Random Forest Classifier — on a dataset, compare their performance, and identify which model gives better accuracy.

**Dataset**

For this task, the Iris dataset from Scikit-learn was used. It is a built-in dataset containing measurements of iris flowers with three classes: Setosa, Versicolor, and Virginica.  
  
Each sample includes the following features:  
- Sepal Length  
- Sepal Width  
- Petal Length  
- Petal Width

**Implementation Steps**

1. Import required libraries: numpy, pandas, sklearn
2. Load the Iris dataset using load\_iris()
3. Split data into training and testing sets (80% / 20%)
4. Train two models: Logistic Regression and Random Forest Classifier
5. Predict results and calculate accuracy for both models
6. Compare the two models’ accuracy scores

**Results**

|  |  |
| --- | --- |
| Model | Accuracy |
| Logistic Regression | 96.67% |
| Random Forest | 100% |

**Conclusion**

Both Logistic Regression and Random Forest performed very well on the Iris dataset. However, Random Forest achieved the highest accuracy (100%), showing better generalization and performance for classification tasks.

**References**

* Scikit-learn Documentation: <https://scikit-learn.org/stable/>
* Kaggle: Intro to Machine Learning