**Task 2 – Predictive Analysis Using Machine Learning**

**Objective:**

To build a machine learning model using classification techniques to predict the species of iris flowers based on sepal and petal measurements.

**Dataset:**

**Iris Dataset**

* Features: sepal\_length, sepal\_width, petal\_length, petal\_width
* Target: species (Setosa, Versicolor, Virginica)

**Process Summary:**

1. **Data Preprocessing:**
   * Checked for null values and data types
   * Dataset found to be clean (no missing values)
   * Label encoding was applied to convert species names into numerical format
2. **Feature Selection:**
   * Used all 4 numerical features as they are relevant for classification
   * Features: sepal and petal dimensions
3. **Model Selection:**
   * Chose **Logistic Regression** (simple yet effective for multi-class classification)
   * Split data into train and test sets (80% train, 20% test)
4. **Model Training & Evaluation:**
   * Accuracy achieved: ~96.67%
   * Evaluation Metrics:
     + **Confusion Matrix**
     + **Classification Report** (Precision, Recall, F1-score)

**Insights:**

* **Petal length and width** are the most influential features in distinguishing species.
* The model performs **very well across all 3 classes**, with high precision and recall.
* Simple classification algorithms like Logistic Regression can still provide strong performance when data is clean and well-structured.

**Conclusion:**

The project successfully demonstrates how to implement predictive analytics using Python and Scikit-learn. The model can accurately predict iris species based on 4 input measurements and can be extended to other classification tasks with similar structure.