FINAL COMBINED REPORT

Internship Domain:

Artificial Intelligence & Machine Learning

TASK – 1

Iris Flower Classification using Machine Learning

Objective:
To classify iris flowers into Setosa, Versicolor, and Virginica species using sepal and petal measurements.

Dataset:
Iris Dataset (scikit-learn)

- 150 samples
- 4 numerical features
- 3 classes

Methodology:

- Loaded dataset using scikit-learn
- Split data into training and testing sets (80:20)
- Applied Logistic Regression classifier
- Evaluated using accuracy score and confusion matrix

Results:

- Accuracy: **100%**
- No misclassification observed

TASK – 2

Spam Mail Detection using Machine Learning

Objective:
To classify text messages as spam or non-spam (ham) using NLP techniques.

Dataset:
SMS Spam Collection Dataset (UCI Repository)

Methodology:

- Loaded text dataset
- Cleaned text (lowercasing, removing symbols)
- Converted text into numeric form using TF-IDF
- Split dataset into training and testing sets
- Applied Naive Bayes classifier
- Evaluated using accuracy, confusion matrix, and F1-score

Results:

- Accuracy ≈ **95–97%**
- High precision and recall for spam detection

Skills Gained

- Data preprocessing
- Machine Learning classification
- Natural Language Processing
- Model evaluation and analysis