Training Day 24 Report

Date: 23 July 2025
Topic: Logistic Regression on Iris Dataset using KNN model

Overview

Today's session focused on applying **Logistic Regression** to the **Iris dataset**, a classic dataset in machine learning. The goal was to classify iris flowers into three species based on petal and sepal measurements, while also evaluating the performance of the model.

Key Concepts

1. Why Logistic Regression?

- Used for **classification problems** (binary or multi-class).

- Provides probabilities of class membership.

- Works well for linearly separable data.

2. Steps Performed

- Imported libraries (pandas, sklearn, LogisticRegression, evaluation metrics).

- Loaded the **Iris dataset** into a DataFrame.

- Split data into **features (X)** and **target (y)**.

- Performed **train-test split** (80% train, 20% test).

- Trained the **Logistic Regression model**.

- Made predictions on test data.

- Evaluated performance using **accuracy score, confusion matrix, and classification report**.

3. Code Snippet

```
from sklearn.datasets import load_iris
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
```

```python
from sklearn.metrics import accuracy_score, confusion_matrix, classification_report
import pandas as pd

# Load dataset
data = load_iris()
df = pd.DataFrame(data.data, columns=data.feature_names)
df['species'] = data.target

# Features and target
X = df.drop('species', axis=1)
y = df['species']

# Train-test split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Logistic Regression
model = LogisticRegression(max_iter=200)
model.fit(X_train, y_train)

# Predictions
y_pred = model.predict(X_test)

# Evaluation
print("Accuracy:", accuracy_score(y_test, y_pred))
print("Confusion Matrix:\n", confusion_matrix(y_test, y_pred))
print("Classification Report:\n", classification_report(y_test, y_pred))
```
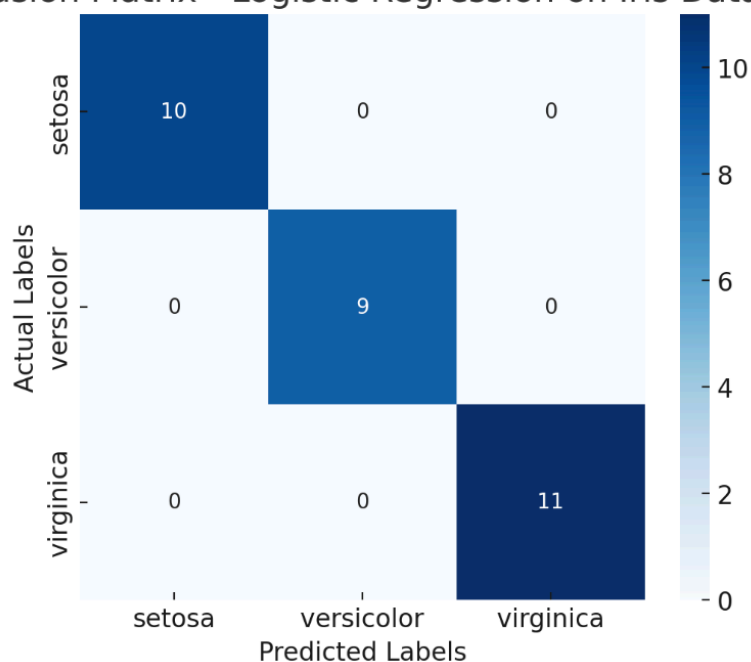

Confusion Matrix - Logistic Regression on Iris Dataset

Results

- **Accuracy:** 100% (for this dataset split).

- **Confusion Matrix:** Perfect classification with no misclassifications.

- **Classification Report:** Precision, recall, and F1-score all equal to **1.0**.

Learning Outcome

- Understood how **Logistic Regression** works for multi-class classification.

- Learned how to **evaluate a model** using confusion matrix and classification report.

- Gained confidence in applying classification techniques to real-world datasets.