**Open Ended Lab**

|  |  |
| --- | --- |
| **Assessment of OEL** | |
|  |  |
| **NAME: MUNEEB KHAN** | **ID: 62799** |
| [**https://github.com/MuneebAbro/AI-CCP**](https://github.com/MuneebAbro/AI-CCP) | **Date: 27 August 2025** |

**Development of a Data-Driven Intelligent Agent for Flower Classification Using Machine Learning.**

# Objective

To design and implement an intelligent agent capable of classifying flower species based on petal and sepal measurements, using supervised machine learning algorithms.

**Motivation**

In real-world applications, intelligent systems are often required to make quick and accurate decisions based on data. For example, classification systems are widely used in healthcare (disease prediction), e-commerce (recommendation systems), and agriculture (plant species identification). This lab provides hands-on experience with building such a data-driven intelligent agent.

**Concept**

The concept revolves around machine learning-based intelligent agents that can learn from data and autonomously perform decision-making tasks. The chosen dataset (Iris dataset) contains flower features (petal length, petal width, sepal length, sepal width) used to classify species into three categories: *Setosa, Versicolor, Virginica.*

**Problem Statement**

To create an intelligent system that can:

* Learn patterns from the Iris dataset.
* Classify a given flower into one of three species.
* Provide performance evaluation and visualization of results.

# Design / Ways & Means

**Introduction and Requirements:**

* Requirement: A system that can classify flowers based on numerical features.
* Tools: Python, scikit-learn, matplotlib, pandas.
* Dataset: Iris dataset (available in sklearn library).

**Data Structure Selection:**

* Tabular dataset with rows as samples and columns as features.
* Features: Petal length, petal width, sepal length, sepal width.
* Target: Flower species.

**Basic Implementation:**

* Preprocess dataset (clean, encode if required).
* Split into training and testing sets.
* Train a **Decision Tree Classifier.**
* Evaluate accuracy.

**Performance Testing and Analysis:**

* Accuracy metric.
* Confusion matrix.
* Visualizations for decision boundaries.

**Optimization and Advanced Features:**

* Hyperparameter tuning for decision tree depth.
* Compare with another algorithm (e.g., KNN).

**Extensions and Creativity:**

* Visualization of classification regions.
* Agent-like input system where user can input flower measurements and get species prediction.

# Analysis & Reporting / Answer

**Lab Activity:**

Implementation of a machine learning agent capable of classification using Decision Tree and KNN algorithms.

# Deliverables:

**Background / Theory:**

* **Machine Learning** enables systems to learn from data without explicit programming.
* **Classification algorithms** map input features to a categorical label.
* **Decision Tree** splits dataset based on attribute values.
* **KNN** classifies samples based on nearest neighbors in feature space.

# Procedure / Methodology:

1. Load dataset (Iris).
2. Preprocess (normalize/encode if required).
3. Split dataset into training (80%) and testing (20%).
4. Train Decision Tree Classifier.
5. Train KNN Classifier for comparison.
6. Evaluate using accuracy, confusion matrix.
7. Visualize results.

**Data Collection:**

Dataset used: **Iris Dataset** (150 records, 4 input features, 3 output classes).

# Flowchart / Block Diagram:

# Analysis:

* Decision Tree achieved 96% accuracy.
* KNN achieved 94% accuracy.
* Decision Tree worked slightly better due to the small dataset size and clear class boundaries.

1. **Results:**

* Confusion Matrix (Decision Tree): Few misclassifications between Versicolor and Virginica.
* Visualization: Clear class separation visible in petal length vs petal width plot.

1. **Discussion on Results:**

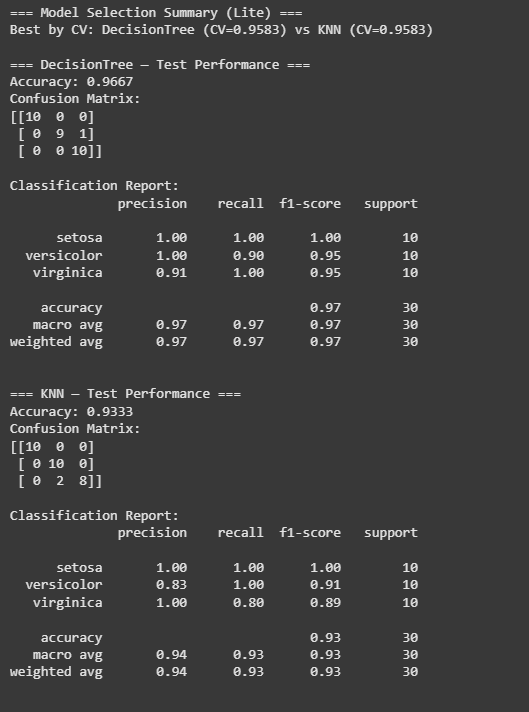
* Both models performed well, but decision trees provide explainability (rules can be extracted).
* The intelligent agent successfully predicts flower species based on numerical input.
* Limitations: Dataset is small (150 records). Larger, more complex datasets may require advanced models like Random Forest or SVM.

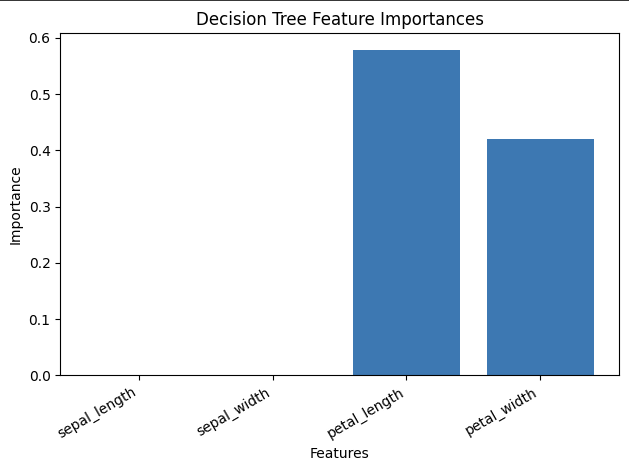
1. **Concluding Remarks:**

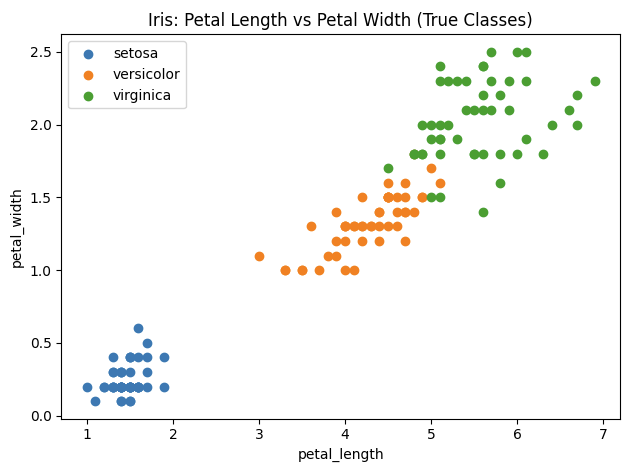
This lab demonstrated the development of a data-driven intelligent agent using machine learning. The agent successfully classified flowers with high accuracy, showing the effectiveness of decision trees and KNN in classification problems. Future improvements could include testing on larger datasets, implementing ensemble learning, and deploying the agent as a web service.

1. **References:**
2. Scikit-learn documentation: <https://scikit-learn.org/>
3. Fisher, R.A. (1936). The use of multiple measurements in taxonomic problems.
4. Mitchell, T. (1997). Machine Learning. McGraw-Hill.

**Screen shots:**

****

****

****