

Problem Statement

Title: *Classifying Iris Flower Species Using K-Nearest Neighbors (KNN)*

Objective

To develop a **K-Nearest Neighbors (KNN) classification model** that predicts the **species of an iris flower** based on its measurable features:

- **Petal length**
 - **Petal width**
-

Background

The Iris dataset is a classic problem in pattern recognition. It consists of three different species of iris flowers: *Setosa*, *Versicolor*, and *Virginica*. KNN is a **non-parametric, instance-based learning algorithm** that classifies new instances based on the majority class among its **k nearest neighbors** in the training data.

Dataset Description

Each sample includes:

- **X₁**: Petal length (cm)
- **X₂**: Petal width (cm)
- **Y**: Species (Setosa, Versicolor, Virginica)

Sample Dataset (Subset)

Sample	Petal Length (cm)	Petal Width (cm)	Species
1	1.4	0.2	Setosa
2	1.3	0.2	Setosa
3	4.7	1.4	Versicolor
4	4.5	1.5	Versicolor
5	5.1	1.9	Virginica
6	5.9	2.1	Virginica

Goals

- Implement a **KNN classification model** using petal measurements
- Use **Euclidean distance** to compute similarity
- Choose an appropriate value of **k**
- Predict the species of new iris samples

- Evaluate the model using:
 - **Confusion Matrix**
 - **Accuracy**
 - **Precision, Recall, F1 Score**
-

Assumptions

- Data points with similar petal dimensions belong to the same class
 - Classes are mutually exclusive and exhaustively defined
 - Features are appropriately scaled or normalized (if needed)
-

Example Task

Predict the species of a flower with:

- Petal length = 4.6 cm
- Petal width = 1.5 cm

Using **k = 3**

Solution
