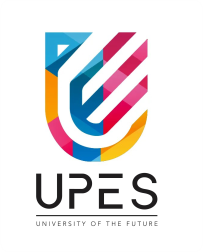
Pattern Recognition

LAB

**Experiment-4**



**NAME : Ribhu Bhushan Tiwari**

**SAP ID : 590011166**

**BATCH : AI / ML (B-2)**

**SUBMITTED TO : Dr. Roohi Sille**

**Unsupervised Learning, k-Means Clustering & Elbow Rule**

**Objective:**

1. To understand the concept of unsupervised learning.
2. To implement k-Means clustering algorithm.
3. To determine the optimal number of clusters using the Elbow Rule.

**Theory:**

1. **Unsupervised Learning:**
   * Unsupervised learning is a type of machine learning where the model is trained on data without labeled responses. The goal is to find hidden patterns or intrinsic structures in the input data.
   * Common techniques include clustering, dimensionality reduction, and association rule learning.
2. **k-Means Clustering:**
   * k-Means is a popular clustering algorithm used to partition data into k clusters.
   * The algorithm works iteratively to assign each data point to one of k clusters based on the nearest mean (centroid).
   * Steps:
     1. Initialize k centroids randomly.
     2. Assign each data point to the nearest centroid.
     3. Recalculate the centroids as the mean of all points in the cluster.
     4. Repeat steps 2 and 3 until convergence (no change in centroids).
3. **Elbow Rule:**
   * The Elbow Rule is a method to determine the optimal number of clusters (k) in k-Means clustering.
   * It involves plotting the within-cluster sum of squares (WCSS) against the number of clusters (k).
   * The "elbow" point in the graph (where the rate of decrease sharply changes) indicates the optimal number of clusters.

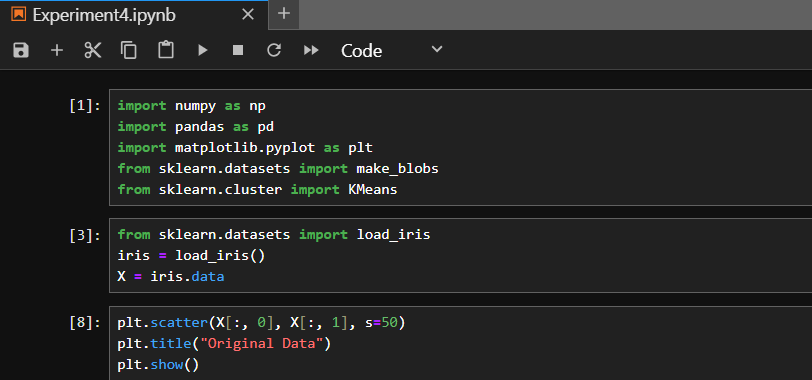
**Tools and Libraries:**

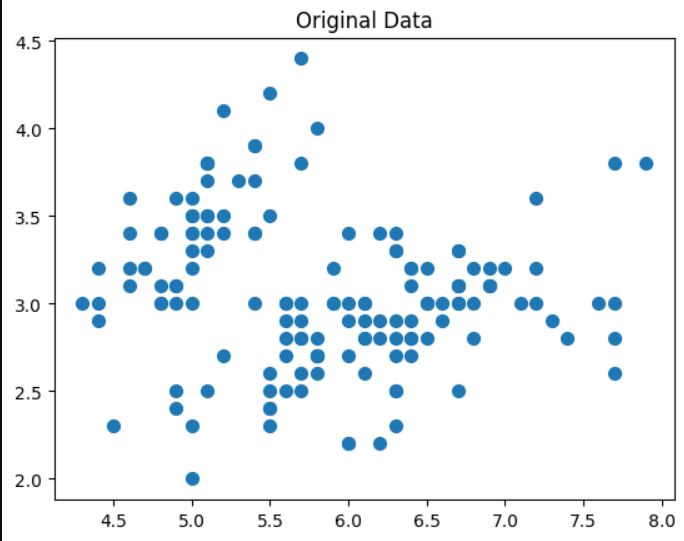
* Python
* Libraries: numpy, pandas, matplotlib, sklearn

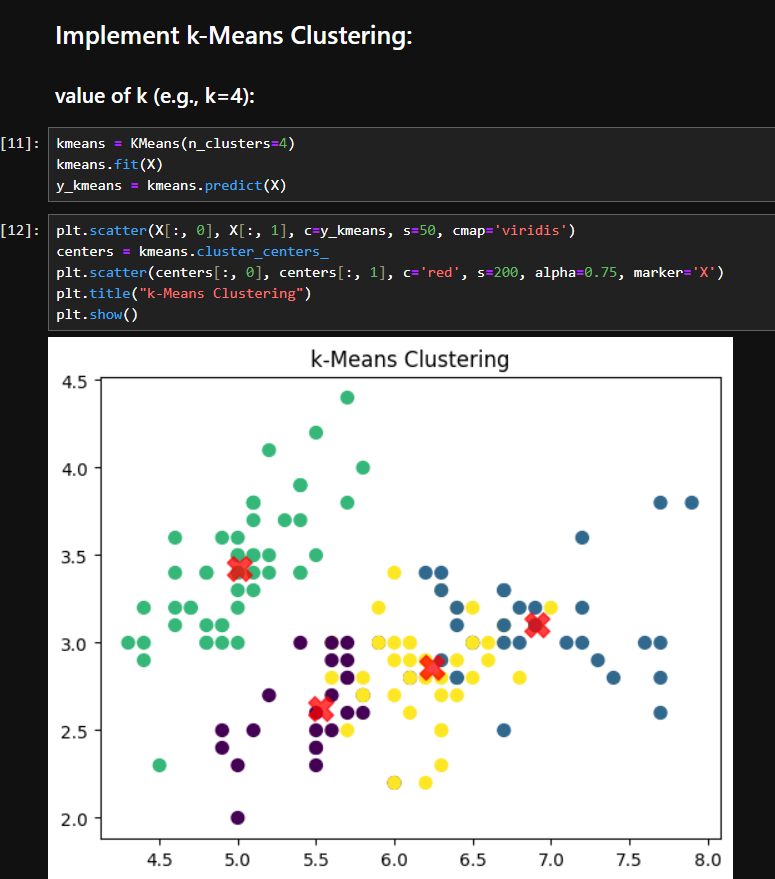
**Dataset:**

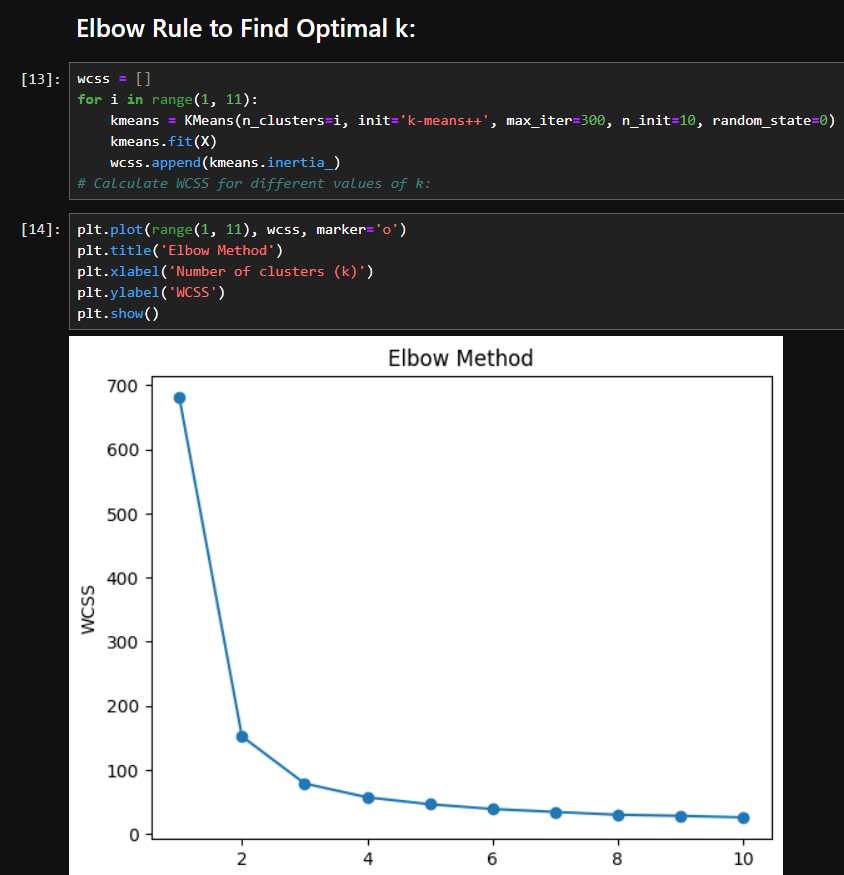
Use a built-in dataset like the Iris dataset

**CODE:**

****

****

****

****

**Observations:**

1. The k-Means algorithm successfully partitions the data into clusters.
2. The Elbow Rule helps in identifying the optimal number of clusters.

**Conclusion:**

* k-Means clustering is an effective unsupervised learning technique for grouping data into clusters.
* The Elbow Rule provides a systematic way to determine the optimal number of clusters, ensuring better model performance.