Lab Assignment 2

Objectives: -

a) Load a dataset containing student marks in 4 subjects. Normalize the data and apply KMeans

clustering to group students based on performance.

Example CSV Dataset:

Student, Math, Physics, Chemistry, Biology

A, 80, 75, 85, 90

B, 60, 65, 70, 72

C, 95, 90, 92, 89

...

Note: Add at least 20 rows in the csv data.

Perform these tasks:

- Read the CSV.

- Normalize the numeric features.

- Apply K-Means (e.g., with k=3).

- Visualize clusters using PCA or t-SNE.

b) Load a dataset of employees with categorical and numerical features,

encode the categorical data and train a decision tree to predict job

satisfaction.

Example CSV Dataset:

ID, Department, Gender, Age, Salary, Satisfaction_Level

1, Sales, Male, 28, 40000, High

2, HR, Female, 34, 35000, Low

...

Note: Add at least 20 rows in the csv data.

Perform these tasks:

- One-hot encode categorical features.

- Normalize numerical ones if needed.

- Train a Decision Tree classifier.

- Print accuracy on test data.