

Lab 1 (due: Jan 24)

MACHINE LEARNING - COSC 4360

Department of Computer Science and Electrical Engineering

Spring 2025

Exercises

Create a **New Project** for every exercise. Take a screenshot of the source code along with its output and place the **source code** and the **screenshot** in a **zipped folder** named **LastNameFirstName_Lab1**

Exercise 1

Download the *wine* dataset from the following Machine Learning repository: [ML Repository](#). Rename *wine.data* to *wine.data.csv*. In *wine.names*, you can read the descriptions of all the attributes.

The dataset contains 13 features (columns 2-14) that contribute to the quality of wine. The dataset contains data for three types, i.e., labels, of wines, identified by the category values: 1, 2, and 3 (column 1). The dataset contains 178 records. Perform classification using **KNN** and, compute the **accuracy** of your model, print the **confusion matrix**, and **predict** to which one of the *three* classes the following *four* wines belong to given the following feature values:

[14.23,1.71,2.43,15.6,127,2.8,3.06,.28,2.29,5.64,1.04,3.92,1065],

[12.64,1.36,2.02,16.8,100,2.02,1.41,.53,.62,5.75,.98,1.59,450],

[12.53,5.51,2.64,25.96,1.79,.6,.63,1.1,5.82,1.69,515],

[13.49,3.59,2.19,19.5,88,1.62,.48,.58,.88,5.7,.81,1.82,580]

Note 1: Use a **test size** of 30% and **K=5**

Note 2: The names of all columns are:

names = ['class', 'Alcohol', 'Malic Acid', 'Ash', 'Acadlinity', 'Magnisium', 'Total Phenols', 'Flavanoids', 'NonFlavanoid Phenols', 'Proanthocyanins', 'Color Intensity', 'Hue', 'OD280/OD315', 'Proline']

Exercise 2

Given the `iris.data.csv`, produce the following two plots as shown in Figs. 1 and 2 below:

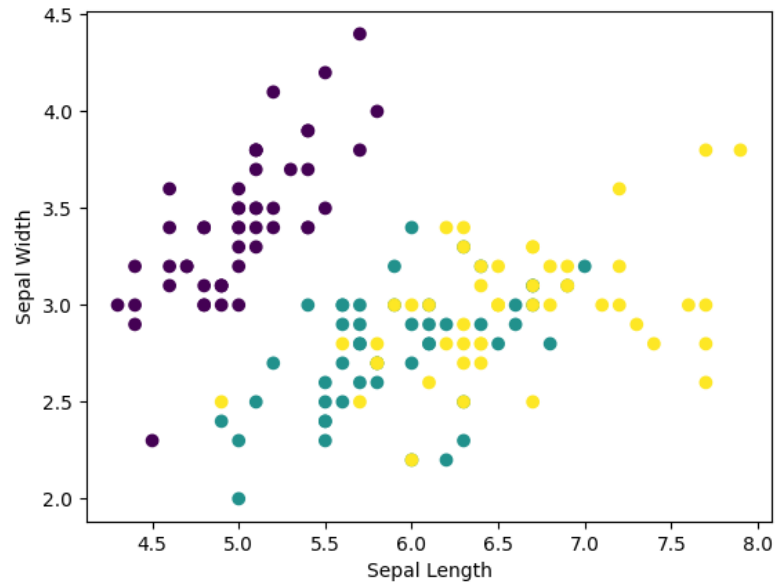


Figure 1: Sepal features

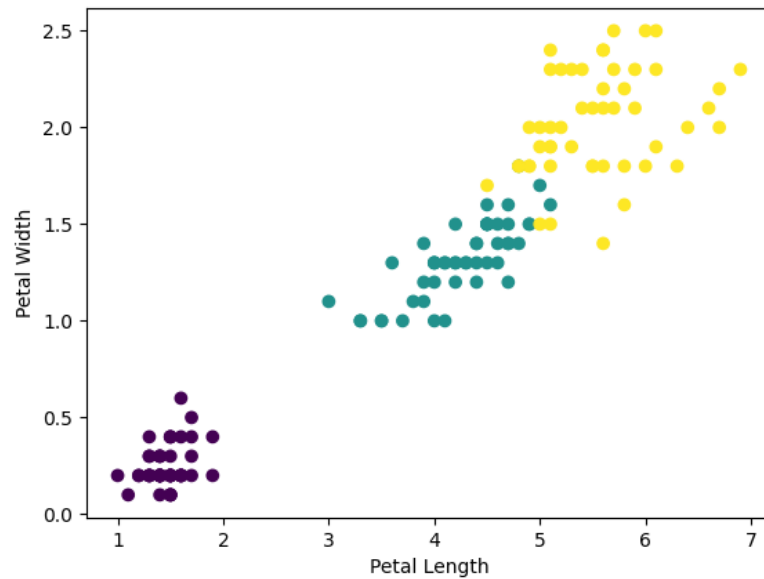


Figure 2: Petal features

Note: Submit through **Canvas**