Assignment 2 due 29th September

This Assignment contains 8 % of the total grade [ Each problem weighs 4 % of total grade.]

Both the problems are 10 points each.

The problems needed to be solved using python.

Do not use any packages beside NumPy, SciPy and Matplotlib for the solution.

Mention to the grader how you want her to take the input in a *readme* file.

You will be submitting all the files required to run your program.

Problem 1 – the evaluator can put any input of their choice. So, test it with all possible cases.

For any questions contact – Irfana Lnu [grader information in Syllabus]

1. Given an array of points where points[i] = [xi, yi] represents a point on the **X-Y** plane and an integer k, return the k closest points to the origin (0, 0).

The distance between two points on the **X-Y** plane is the Euclidean distance (i.e., √(x1 - x2)2 + (y1 - y2)2).

You may return the answer in **any order**. The answer is **guaranteed** to be **unique** (except for the order that it is in).

**Example 1:**

A graphing a cross with points

Description automatically generated with medium confidence

**Input:** points = [[1,3],[-2,2]], k = 1

**Output:** [[-2,2]]

**Explanation:**

The distance between (1, 3) and the origin is sqrt(10).

The distance between (-2, 2) and the origin is sqrt(8).

Since sqrt(8) < sqrt(10), (-2, 2) is closer to the origin.

We only want the closest k = 1 points from the origin, so the answer is just [[-2,2]].

**Example 2:**

**Input:** points = [[3,3],[5,-1],[-2,4]], k = 2

**Output:** [[3,3],[-2,4]]

**Explanation:** The answer [[-2,4],[3,3]] would also be accepted.

 These are examples. The problem is user input based. The grader can test with any input of their choice.

**Constraints:**

* 1 <= k <= points.length <= 104
* -104 <= xi, yi <= 104

1. Using the iris data set of assignment 1 [[**https://archive.ics.uci.edu/dataset/53/iris**](https://archive.ics.uci.edu/dataset/53/iris)**,**

using python convert **.data** files into **.csv**.] split it into

* 1. 80% train and 20% test data
  2. 70% train and 30% test data
  3. Compare the accuracy, specificity and sensitivity of **a** and **b** and with the ROC curve mention which is a better model. The grader should be able to generate the curve and the stats after running your program.