UNITED INTERNATIONAL UNIVERSITY

Department of Computer Science and Engineering (CSE)

Evaluation 2

40 min

Course : DSA I Lab	**Course Code: CSE 2216**
Trimester & Year: Summer 24	**Section N**
Name :	**ID:**

SN	Questions	Marks
1	**Instructions:**	10+5= 15

- Launch VS Code or any other IDE and ensure that you can import the two libraries - **numpy** and **pandas**
- If you are **unable** to import any of the two libraries, type the following command in the terminal
  - pip install numpy
  - pip install pandas
- Download the previously given code done in class.
- You **MUST NOT** use any libraries other than numpy and pandas
- Download the **dataset1 and dataset2**

Task 1:

You have been given a dataset (dataset1) that shows the predicted marks and performance levels of several students based on their study hours, attendance percentage, number of assignments completed, and number of research papers studied. Your task is to create a function that takes a new student's data as input (including study hours, attendance percentage, number of assignments, and number of research papers) and applies the K-Nearest Neighbors (KNN) algorithm with k=3 to predict the performance level of this student using the provided dataset.

You must ensure the following:

- Ensure that you use only the numpy and pandas libraries for this task.
- Define the distance function based on the dataset
- Do not use any other libraries - code everything from scratch as shown in last class
- You can use the previously completed code

Sample Function Call	Sample Output

# study hours, attendance, assignments , research papers 	
x_test = [6, 90, 5, 4]
y=applyKNN()
print(f‘y={y}’)
	y= 815

Task 2:

After developing the KNN function in Task 1, you will test its performance on another dataset (**dataset2**). To evaluate its accuracy, you will measure the error of each prediction and then calculate the average error across all predictions. This process will help you understand how well your KNN algorithm performs on new, unseen data.

Sample Output

```
Predicted Values
[535.0, 248.33333333333334, 230.0, 863.3333333333334]
Actual values
[600.0, 405.0, 337.5, 802.5]
Errors
[10.833333333333334, 38.68312757201646, 31.851851851851855, 7.580477673935622]
Average error % = 22.237197607784317
```