National University
of Computer & Emerging Sciences-Islamabad
Chiniot-Faisalabad Campus

Dr. Hashim Yasin
Department of AI & DS.

AI3002 –Machine Learning
Assignment No. 4

Assignment Submission Guidelines:

1. Submit your assignment in **soft form (Code + Report)** within the due date and time. Soft form does not mean submitting photos of the hardcopy. Late submissions will result in a deduction of marks.
2. The **report** must include a discussion, comments, and a conclusion about your solution. Submitting without a report will result in a loss of full marks.
3. Name the zip or other folder/file that you submit using the following format: **ML_A4_RollNo_FirstName**.
4. Ensure that you solve each task of the assignment on your own.
5. You are allowed to do your assignment in **groups of a maximum of two members**.
6. There is no restriction on the programming language used for the tasks.
7. For programming tasks, you are NOT allowed to use any built-in functions or libraries for specific tasks.

Question 1: K-means, K-medoids, and K-median Clustering

Using the dataset provided in the "dataset" folder, which contains four features (A1, A2, A3, and A4), perform the following tasks:

1. **K-means Clustering:** Implement the K-means clustering algorithm, where the number of clusters K is defined by the user (e.g., K could be 2, 3, or 4).

2. **K-medoids Clustering:** Implement the K-medoids clustering algorithm, allowing the user to define the number of clusters K (e.g., 2, 3, or 4).

3. **K-median Clustering:** Implement the K-median clustering algorithm with a user-defined K value (e.g., 2, 3, or 4).

4. **Cluster Visualization:** Create scatter plots for each clustering method, coloring the data points based on their assigned clusters for K-means, K-medoids, and K-median clustering.

5. Discuss the advantages and disadvantages of each clustering algorithm used above.

Question 2: Hierarchical Clustering

Using the dataset in the "dataset" folder with features A1, A2, A3, and A4, implement the following tasks:

1. Single Linkage Agglomerative Clustering

2. Complete Linkage Agglomerative Clustering

3. Average Linkage Agglomerative Clustering

Draw a dendrogram and a Venn diagram for each clustering method above.