# CS 436/536 Assignment-5

**Description:**

Un-supervised Learning. Please refer provided notebook

1. In the provided notebook there is a dataset.
2. Remove the labels from the dataset to make the number of classes unknown and plot the unlabeled data again.
3. Apply K-means clustering algorithm on the created unlabeled dataset.
4. Implement your own k-means clustering algorithm and do not use any built-in library.
5. Vary K from 2-9 and evaluate the k-means algorithm's performance based on any Quantitative evaluation metric of k-means of your choice.
6. Plot the K vs performance.
7. Label the dataset again with k of your choice and plot it.

**Submission:**

1. Due date is midnight of 13 April 2023. The total points will be reduced by 5% for each day after the due date.
2. Submission in a single PDF file. Include all the code with comments, plots, and summary.
3. Generate the unlabeled dataset from labeled and plot. [10 Points]
4. Apply K-means clustering algorithm on the created unlabeled dataset [40 Points]
5. Vary K from 2-9 and evaluate the k-means algorithm's performance based on any Quantitative evaluation metric of k-means of your choice and plot it. [20 Points]
6. Give convincing reasoning why would you choose a certain value of K for this problem. [20 Points]
7. Label the dataset again with your preferred k and plot in different color in same graph. [10 Points]

**Note:**

1. It is important that everyone submits their entire notebook code in PDF format as well merged to the report. If this is not done points will be deducted.
2. Feel free to provide links to your notebook in your report as well.
3. Review and follow these Watson College Academic Honesty policies that spell out the consequences of academic dishonesty.
4. Do not copy/give code from/to others. If plagiarism is found, both will receive zero points.
5. You can submit multiple times before the due date, only the last submission will be graded.

**Important Notes:**

Make a report first answering all the questions individually. For example, if a question is asking you to show a plot, copy your plot from your notebook and paste it in your report with the question. If a question is asking you to report the final accuracy, please put the values in the report with the question.

After the report is done, then merge a pdf of your notebook to the bottom of your report. Please submit only a single PDF file.