CS 564: Foundations of Machine Learning
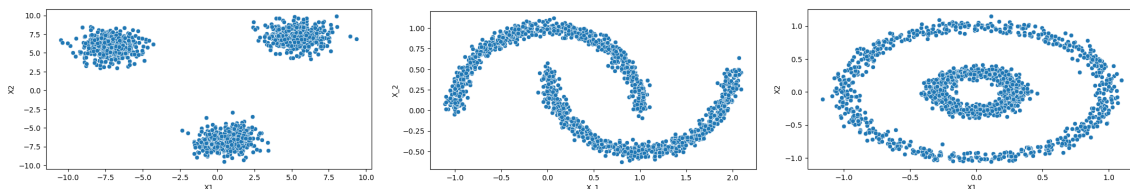## Assignment 2

Deadline: 13 September 2022

- Markings will be based on the correctness and soundness of the outputs.
- Marks will be deducted in case of plagiarism.
- Proper indentation and appropriate comments (if necessary) are mandatory.
- Use of frameworks like scikit-learn etc is allowed.
- *All benchmarks(accuracy etc), answers to questions and supporting examples should be added in a separate file with the name 'report'.*
- *All code needs to be submitted in '.py' format.* Even if you code it in '.ipynb' format, download it in '.py' format and then submit
- You should zip all the required files and name the zip file as:
  - <roll_no>_assignment_<#>.zip, eg. 1501cs11_assignment_01.zip.
- Upload your assignment (the zip file) in the following link:
  - https://www.dropbox.com/request/Y7uWRDGgJj2uKkiOYDGi

Problem Statement:
- The assignment targets to implement DBScan algorithms to cluster the 3 datasets with blob, moon and circle structures. Each csv file consists of two columns in the datasets.
  - https://www.dropbox.com/scl/fo/9jbpw1ah58jcvva5bzwdl/h?dl=0&rlkey=4ao0wdd9ipyvwuhpp0edvqic2

Implementation Steps:
- Design and implement DBSCAN algorithm to cluster the 3 datasets mentioned. Assume hyperparameters wherever necessary.
- Use Silhouette index to evaluate the clustering quality.
- Run dataset for the K-means algorithm implemented in assignment 1. Set the num_clusters for K-means to the no. of clusters from DBSCAN. Compare it to DBSCAN in terms of cluster visualization and Silhouette index score.

Documents to submit:
- Model code
- Silhouette index
- Visualization of clusters for DBSCAN and K-means clustering

For any queries regarding this assignment contact:

Abhisek Tiwari (abhisektiwari2014@gmail.com),

Ratnesh Kumar Joshi (ratneshkr.joshi@gmail.com), and

Ramakrishna Appicharla (ramakrishnaappicharla@gmail.com)