Lewis University
Department of Engineering, Computing, and Mathematical Sciences

DATA 55100: Unsupervised Machine Learning

Assignment #3

Submission: Submit an electronic copy of a short report (2-5 pages) in Blackboard.

Grading: Your report is the most important part of this assignment. You need to program up the assignment, but your presentation, description and analysis is what I will grade!! In addition, **don't use any built-in packages** in MATLAB and Python. Instead **develop your own code** for all the algorithms and techniques we are going to study.
In addition, you are allowed to use Jupyter IPython notebooks for analysis with documentation as a substitute for a Word document when writing the report.

Description:
In general, the goal of this small assignment is:
1. Apply the Basic Sequential Clustering Algorithm (BSCA) on the results you obtained in week 2.

Specifically, you are to:
1. Apply BSCA algorithm on the dataset(s) you chose for week 2 assignment. Specifically, apply BSCA on the new 2-D and 1-D representation of the datasets after applying PCA and LDA respectively (from week 2). At the end analyze your findings.
2. Experiment BSCA using three different values for the threshold of dissimilarity (α) and three different values for the maximum number of clusters (M).

Your report should contain sections on:
1. The technical description of all techniques utilized
2. The design of the algorithms (pseudo-code, flowcharts, or some other structured descriptive means),
3. The results of the algorithms
4. An analysis of the results, i.e., did you obtain what you expected? Were there any surprises? What conclusions can you draw from the experiments? etc.
5. Well documented, structured, modular program listings.