Lewis University
Department of Engineering, Computing, and Mathematical Sciences

DATA 55100: Unsupervised Machine Learning

Assignment #6

Submission: Submit an electronic copy of a short report (2-4 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 cluster validity indices you will choose in this assignment.
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 assignment is:
1. Validate the **clustering findings** you obtained in assignment #5.

Specifically, you are to:
1. Implement three cluster validity indices (**one must be the CS index**).
2. Taking into consideration the results you obtained after implementing VAT, iVAT (as a pre-clustering technique) and FCM (as a clustering technique), compare the results you will get after applying the validity indices with the previous results of VAT (iVAT) and FCM. What do you think the best choice for the number of clusters (C) will be for each dataset you've analyzed?

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.