Lewis University
Department of Engineering, Computing, and Mathematical Sciences

DATA 55100: Unsupervised Machine Learning

Assignment #4

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 Visual Assessment of Clustering Tendency (VAT) and iVAT algorithms on some datasets.

Specifically, you are to:
1. Produce VAT and iVAT images and analyze the results.

I have provided you with a ZIP called "Clustering.zip" that contains:
1. 2D data sets (as PGM images).
2. MakeDataSet.m, which creates a two dimensional data set from a PGM file.
3. LoadDataSets.m, which loads all data sets in the folder into a cell array.
4. pdist2.m, which computes the pair-wise distance between all vectors in a data set.
5. VAT.m and iVAT.m, which computes the VAT 'image' from the result of pdist2.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.