Name _____

Colorado State University ENGR 571, Analytics in Systems Engineering, Professor Simske

Project #2, Assigned 8 April 2020, Due 13 May 2020 (100 points)

Please select a data-related project from one of the following:

1. **Module 10 classification problem re-done with a different f(z)**
2. **System model for a data set**
3. **Image classification challenge from Module 11, Page 1:** The image classification training and testing XML ZIP files and results are posted on Module 11, Page 1 (scroll down, please)
4. **Data mining of a work-relevant data set**
5. **Unsupervised classification**
6. **Classification model of Module 10 applied to your own data set (you could even apply this to the data set for Project 1, though you may not be happy with the classification accuracy!)**

The final project will be a written document—WORD or PDF format preferred. Please run the topic past me first, if you have not already done so. Incorporate this page as the first page in your report, please.

Sections expected are:

1. **Problem Statement (10 points)**
2. **Experimental Approach (20 points)**
3. **Results (30 points)**
4. **Discussion (30 points)**
5. **Learning and Follow Up (10 points)**

I am not expecting a lengthy report. The key element of the assignment is to apply data science techniques and processes you've learned in the course to a real-world problem. If you have no such problem, please use one of the two course-provided data sets (email logs or image classification).

Questions? Steve.Simske@Colostate.edu