Capstone Project 1 Proposal

Trisomy 21, aka Down Syndrome, includes among its constellation of symptoms, intellectual disability and learning problems. The problem: identify proteins whose expression in mouse brains is specific to Down Syndrome and the experience of a learning task, particularly for interactions between disease and learning conditions.

The data for this project comes from a study investigating the drug memantine in a mouse model of Down Syndrome. Memantine is a NMDA receptor antagonist used to treat severe Alzheimer's. In subsequent clinical studies, it was not found to be helpful in older adults with Down Syndrome. Identifying protein expression differences that identify differences in learning in Down Syndrome brains is an early step in identifying potential targets for treatments for the cognitive effects of trisomy 21. A pharmaceuticals company or academic research lab that is doing basic research aimed at finding potential new treatments for intellectual disabilities can use this to identify new candidate drugs to test.

The data is from the UCI Machine Learning Repository: http://archive.ics.uci.edu/ml/datasets/Mice+Protein+Expression. Being from the UCI collection, it is freely available and fairly clean. It consists of expression data for 77 proteins or protein modifications (e.g. histone methylation) found in the cell nuclei from mouse brains from 8 conditions: wildtype or a mouse model of Down syndrome, with or without a learning task (context fear conditioning), and memantine or placebo. There are 7-10 mice per group and 15 independent measurements per mouse.

The 8 relevant groups are labeled, so this is a supervised classification problem. The deliverables will include the code and a slide deck of my findings.