**COSC 3337 – Final Exam Review**

These are topics you should study and be prepared to have questions over in the final exam. Some questions require calculations, while others require answering questions about an algorithm, it’s assumptions, and/or how it works. Understanding the concepts is key.

* Entropy
  + How to calculate for a given sample of data
  + How to calculate information gain for a decision tree split
* Clustering
  + Understand DBSCAN Method
    - What are min\_samples and eps (epsilon) used for?
  + Know Euclidean and Manhattan Distance
  + Understand Nearest Neighbors Clustering
  + Hierarchical Clustering
    - Difference between Agglomerative and Divisive HC
    - Be able to create dendrogram / do clustering
  + Understand K-Means Clustering
  + Be able to calculate silhouette coefficient for clusters
* Support Vector Machines (SVM)
  + Understand general principles (pros / cons)
  + What parameters are used by SVM? What do they do?
* Neural Networks
  + Understand general structure (pros / cons)
  + How to compute Neural Network
* Association Analysis
  + A priori
    - Understand algorithm
    - What is min\_support (minimum support)?
    - What is confidence?
* Coding
  + How to create plots?
  + Parameter optimization