COMP 4710 – Introduction to Data Mining

Calendar Description: Introduction to data mining concepts and their applications.
Prerequisite: COMP 3380 and [one of STAT 1150, STAT 2000 (B), STAT 2001 (B), STAT 2220, or PHYS 2496].

Outline

1) Review: Database concepts and usage (1 week)
2) Discovery of frequent patterns (2 weeks)
3) Formation of interesting rules (2 weeks)
4) Analysis of sequential data and time series (2 weeks)
5) Clustering and unsupervised learning (2 weeks)
6) Classification and supervised learning (2 weeks)
7) Detection of anomalies (2 weeks)
8) Other topics (if time permits):
   Incremental mining, mining from data streams, Data visualization, Real-life applications of data mining, a brief introduction on distributed data mining, parallel data mining, and machine learning

Text: Han & Kamber, *Data Mining: Concepts and Techniques,* Morgan Kaufmann, 2006
Tan et al., *Introduction to Data Mining*, Addison-Wesley, 2006
Optional References: Hand et al., *Principles of Data Mining,* MIT Press, 2001
Kantardzic, *Data Mining: Concepts, Models, Methods, and Algorithms*, Wiley, 2003
Witten & Frank, *Data Mining: Practical Machine Learning Tools and Techniques,* Morgan Kaufmann, 2005