CMPT 741 - G100 Data Mining

Instructor(s): Ke Wang

Calendar Objective/Description:
Data Mining

Instructor's Objectives:
Data mining aims to extract useful patterns, trends and previously unknown knowledge from raw data for decision support. This course has two focuses: basic concepts and techniques, and recent technologies and developments in dealing with very large data sets. For the first focus, we will study the classic data mining techniques including association, classification, and clustering; for the second focus, we will study the dominant software systems and algorithms for coping with Big Data. Topics include finding similar items, link analysis, recommendation algorithms, data privacy and security. The course will involve assignments/projects, one midterm and final exam.

Prerequisites:
see go.sfu.ca

Topics:
- 1. Introduction
- 2. Association Rule Mining
- 3. Classification and Supervised Learning
- 4. Clustering and Unsupervised Learning
- 5. Finding Similar Items
- 6. Link Analysis
- 7. Recommendation Systems
- 8. Data Privacy and Security

Grading:
Assignments/Projects (40%), Midterm (20%), and Final exam (40%)

Required Books:
Introduction to Data Mining 2nd Edition, Pang-Ning Tan, Addison Wesley, 9780133128901, Available online
Lecture notes: a combination of the notes provided by the authors in item 1, the slides of the course “CS345A: data mining” at Stanford University, and the slides of the instructor, will be available online to enrolled students.
Data Mining: Concepts and Techniques, 3rd Edition, Han, Kamber, Pei, Morgan Kaufmann, 22 Jun 2011, 9780123814791, Available online

Academic Honesty Statement:
Academic honesty plays a key role in our efforts to maintain a high standard of academic excellence and integrity. Students are advised that ALL acts of intellectual dishonesty will be handled in accordance with the SFU Academic Honesty and Student Conduct Policies (http://www.sfu.ca/policies/gazette/student.html).