CMPT 843 - G100 Database and Knowledge-base Systems

Instructor(s): Jian Pei

Calendar Objective/Description:
An advanced course on database systems which focuses on data mining and data warehousing, including their principles, designs, implementations, and applications. It may cover some additional topics on advanced database system concepts, including deductive and object-oriented database systems, spatial and multimedia databases, and database-oriented Web technology.

Instructor’s Objectives:
The purpose of this graduate course is twofold: broadening graduate students’ knowledge and understanding of the current frontiers of data analytics and management research, and teaching data analytics and data mining research methodologies and skills.

In this semester, we will focus on computational statistics and applications. Specifically, we will cover some fundamental and useful ideas and methods, including sampling, EM optimization methods, simulation, bootstrapping, and density estimation, as well as their programming implementation.

Since it is an advanced graduate course, sufficient preparation and interest in data analytics (i.e., database systems and data mining) and solid undergraduate entry level statistics are assumed. The course itself uses mathematics and probability heavily.

Prerequisites:
None

Topics:
- Sampling
- Optimization and simulation
- EM methods
- Bootstrapping
- Density estimation
- MC
- MCMC
- Distributed optimization and data mining

Grading:
Grading will be announced in the first week of the class. Evaluation will be based on assignments, exams and projects.

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).