INSE 691D
Security Implications of Data Mining
Course Outline - Summer 2008 (2nd Term)

Concordia Institute for Information Systems Engineering

Instructor	**Benjamin Fung** – Office: EV7.649 E-mail: fung@ciise.concordia.ca
Lectures	*Monday and Wednesday 6:30pm-9:00pm*
Office Hours	*Monday 4:30pm-6:20pm*

Objectives

Data Mining is a collection of techniques for discovering hidden knowledge in the rapidly growing data in governments, businesses, sciences, internet, and other information sources. Many applications of data mining, however, pose security and privacy threats to the general public. This course studies the security issues caused by the advancement of data mining technologies. Students will develop skills that can be applied to advancing commerce, government, and science through the exploration of the concepts and techniques of data mining and practical exercises. They will also study secure and privacy-preserving methods to prevent potential threats caused by data mining. This course prepares students to be potential researchers and practitioners in the areas of data mining and information security.

In general, the course will include discussions and explanations of the following subjects: Security implications of data mining and its applications, current privacy legislation and national security policies. Security and privacy threats caused by current data mining techniques. Risks and challenges in emerging data mining applications: text mining, web mining, social networks mining, RFID workflow. Attacks and prevention methods: web privacy attacks, data mining-based intrusion detection, privacy-preserving data publishing.

Required Textbook

Data Mining: Concepts and Techniques, Second Edition, by Jiawei Han and Micheline Kamber, Morgan Kaufmann, 2006. ISBN 1-55860-901-6

Background

Prerequisite: COMP 5531 or permission from instructor. The students are expected to have some basic knowledge of databases.

Grading

The grading of the course is as follows:

Midterm: TBD%

Participation: TBD%

Term Project: TBD%

The term project can be in a form of a research paper, a survey paper, and a programming project + a technical report. The topic has to be related to data mining and approved by the instructor. Some sample topics will be given.

Important: There is no substitution to a missing midterm, so make sure that you write the midterm at the scheduled time.

There are no pre-set cutoff points for the final grades; the cutoff points will be decided based on an assessment of difficulty level, class performance, and fairness. That is, there is no definite rule for translation of number grades to letter grades.

Website

The webpage for the course is http://www.ciise.concordia.ca/~fung/INSE691D/. The webpage will contain announcements related to the class, pointers to documents, class notes, project descriptoin, etc. The user-name and password required to access this page will be provided in class.