

Syllabus

CSRU-4631-001 Data Mining

Department of Computer and Information Sciences

Fordham University

Spring 2005

Description of Course:

This course will introduce popular data mining methods for extracting knowledge from data. It will balance theory and practice—the principles of data mining methods will be discussed, but students will also acquire hands-on experience using state-of-the-art software to develop data mining solutions to scientific and business problems. Topics and related methods discussed in the class include: data warehousing and OLAP, data preprocessing, association mining, classification and prediction, cluster analysis, and mining complex data types. Textbook readings will be supplemented with current articles on data mining technology and applications.

Instructor:

- Dr. Gary Weiss
- **Office:** JMH 328A
- **Phone:** 718-817-0785
- **Email:** gweiss@cis.fordham.edu
- **Office Hours:** Monday, Thursday 11:00-11:30, 1:30-2:30, or by appointment

Objectives:

To develop familiarity with data mining techniques and be able to apply them to real-world problems.

Outcomes:

A student who successfully completes this course will be able to:

- Understand basic data mining techniques, how to apply them, and when they are applicable
- Be able to utilize a state-of-the-art commercial data mining package
- Be able to apply data mining techniques to solve problems

Textbook:

Data Mining: Concept and Techniques by Jiawei Han and Micheline Kamber

There will also be supplemental readings assigned during the semester.

Attendance:

It is important to attend every class session. Please notify me in advance if you must miss a class either personally, or through email or a phone call. Three or more unexplained absences will result in a lowering of the final grade.

Class Participation:

Class participation is an essential part of this class. Students are expected to complete the assigned readings by the specified class date and are expected to actively participate in classroom discussions. Perfect attendance does not ensure a good class participation grade. Students will also occasionally be asked to present some material in class and this will also factor into the class participation grade.

Homework/Quizzes:

Homework assignments are an important part of the class, and should be completed on time. Late assignments will be penalized. Homework assignments may take several forms: problem sets, short write-ups of supplemental readings, and in-class presentations. Some homework assignments will require the use of SAS Enterprise Miner, the sophisticated data mining package purchased by the CIS department for your use. There may be several quizzes during the semester to ensure that the readings are completed on time.

Final Course Project:

In addition to homework assignments, there will be an in-depth course project, due at the end of the term. This project can be done individually, or in groups of two. The project will differ for each individual/group, based on their interests and ideas. Each project must be approved by the instructor. A typical project will investigate one issue/hypothesis in data mining, most likely using our data mining package. The description and results of the project must be written up in a report.

Exams:

There will be 1 exam during the semester and a final exam.

Grading:

The percentages given below are guidelines for both the student and instructor and may be changed as needed to reflect circumstances in the course.

Homework/Quizzes	25%
Class Participation	15%
Exam	15%
Final Project	25%
Final Exam	20%

Blackboard:

This course will make extensive use of blackboard and all assignments—and many supplemental readings—will be located there.

Academic Honesty:

All work produce in this course should be your own unless I specifically specify otherwise. Violations of this policy will be handled in accordance with university policy which can include automatic failure of the assignment and/or failure of the course. In situations where collaboration is permitted or required you should be careful to cite any individual who provided assistance and is not already credited on the work.