

CS499 DATA MINING

General Information:

- College: College of Engineering and Natural Science
- Course: CS499 – Data Mining (3 credit hours)
- Semester: Spring 2007
- Time & Location: Tues, Thurs, 11:10AM-12:25PM, Engineering Bldg 69, Rm 314
- Instructor: Dan Li (Dan.Li@nau.edu)
- Office: Engineering Bldg 69, Rm 255
- Office Hours: Mondays 9am – 12am; Thursdays 1pm – 3pm (or by appointment)
- Course Website: http://www.cens.nau.edu/~dl259/teaching/cs499_spring2007/

Course Prerequisites: CS249 (Data Structures and Algorithms)

Course Descriptions: Data mining is the process of automatic discovery of patterns, changes, associations, and anomalies in massive databases. This course will provide an introduction to the main topics in data mining and knowledge discovery, including: data preparation for knowledge discovery, design and implementation of data warehouse and OLAP systems, data cube computation and data generation, association mining, classification, and cluster analysis.

Learning Outcomes: Successful completion of this course will provide a student with the necessary skills to design basic data mining algorithms to solve a variety of real-world applications.

Course Structure/Approach: This course presents material during lectures and through assigned reading and homework that are expected to be completed outside of class.

Textbook and Required Materials: *Data Mining Concepts and Techniques*, Jiawei Han and Micheline Kamber, ISBN: 1558609016.

Course Outline: The schedule is tentative and subject to change depending upon the progress of the class.

Week	Topic
1. (Jan 15- Jan 19)	Introduction
2. (Jan 22- Jan 26)	Data Preprocessing
3. (Jan 29- Feb 2)	Data Preprocessing, DW & OLAP
4. (Feb 5- Feb 9)	DW & OLAP
5. (Feb 12- Feb 16)	Data Cube Computation, Midterm I
6. (Feb 19- Feb 23)	Data Generation
7. (Feb 26- Mar 2)	Frequent Patterns
8. (Mar 5- Mar 9)	Frequent Patterns
9. (Mar 12- Mar 16)	Classification & Prediction
10. (Mar 19 – Mar 23)	Spring Break
11. (Mar 26 - Mar 30)	Classification & Prediction, Midterm II
12. (Apr 2 – Apr 6)	Classification & Prediction
13. (Apr 9 – Apr 13)	Cluster Analysis
14. (Apr 16 – Apr 20)	Cluster Analysis
15. (Apr 23 – Apr 27)	Cluster Analysis, Midterm III
16. (Apr 30 – May 4)	Project Presentations
17. (May 7 – May 11)	Final Exam, Project Report Due

Assessment of Student Learning Outcomes:

- **Methods of Assessment:**

Assessing students' achievement of this course's learning outcomes is done using homework, in-class participation, examinations, and project assignments.

(1) Homework: Homework will be collected and checked for both completion as well as content.

(2) Examinations: You will have to work efficiently to complete exams in the time allotted. The exams will be very similar to the homework problems, hence, if one does not do the homework the likelihood of successfully writing the examinations is greatly diminished.

(3) In-class Participation: In-class participation consists of presentation, answering questions, and in-class quizzes.

(4) Project: You will do team project which will be assessed through project presentation and project reports.

- **Timeline for Assessment**

Homework is typically assigned tri-weekly throughout the semester. The mid-term exams will be given in week 5, week 11, and week 15. The final exam will be given during finals week.

Grading System:

	Numbers	Total Weight
Assignments	4	20%
Project	1	20%
Midterm Exam	3	20%
Final Exam	1	20%
In-class Participation	N/A	20%

Grading: A=90%-100%, B=80%-89%, C=70%-79%, D=60%-69%, F<60%.

Course Policies:

- **Makeup Tests**

There are NO make-up tests without prior consent of the instructor.

- **Attendance**

Attending class is MANDATORY; tardiness and unexcused absences are considered lack of in-class participation and will be reflected in your grade.

- **Homework Late Policy**

You are given TWO "late days" that can be used towards homework submissions during the semester. A late day is defined as exactly 24 hours, so if an assignment is due at 11:10AM on Tuesday, you could submit the assignment by 11:10AM on that Wednesday by using up one of your late day.

- **Plagiarism and Cheating**

One word: DON'T. Cheating on tests or assignments will result in an immediate failure in the course. Cheating is intentionally claiming credit for work or knowledge that is not your own. Cheating is also intentionally making it possible for others to claim credit for work or knowledge that is not their own. In the event of a cheating incident the student receives an F in the course and a formal record of the incident is put into the student's file.

Moreover, you'd be amazed how easy it is to detect plagiarism or cheating. Cheaters don't spend tremendous amounts of time masking their copy, because that defeats the purpose and it would be simpler to do the homework themselves. Invariably, therefore, they get caught.

University Policies:

- **Safe Working and Learning Environment**

<http://oak.ucc.nau.edu/dam1/Safe%20Policy.htm>

- **Students with Disabilities**

<http://www2.nau.edu/dss/>

- **Institutional Review Board**

<http://www4.nau.edu/ovp/irb/index.htm>

- **Academic Integrity Guidance**

<http://www.nau.edu/library/information/guides/plagiarism.html>