

CSC 474 / 674 Principles of Data Mining

Course and Contact Information

Credits: 3

Meeting Times: MW 5:30pm-6:45pm

Location: Petty Building 136

Instructor: Lixin Fu

Office: Petty 162

Tel.: (336) 402-9601

E-mail: lfu@uncg.edu

Office hours: MW 1:30pm -- 3:00pm or by appointment

Course Objectives

This course gives a comprehensive and state-of-the-art coverage of data mining concepts and techniques for uncovering interesting data patterns hidden in large data sets. The major objectives and student learning outcomes are

- students understand basic concepts of data warehousing (DW), OLAP, and data mining (DM)
- students will learn about the classifications, architectures, and models of DW and DM systems; challenges and issues; applications and commercial potentials
- students should be able to understand how major DM (association rules, classification, and clustering) algorithms work
- graduate students are prepared for DM research or industry through writing term papers on topics or implementing DM algorithms

Textbook and Readings

Required textbook:

Jiawei Han, Micheline Kamber, Jian Pei, Data Mining: Concepts and Techniques, Third Edition, Morgan Kaufman Publishers, ISBN 978-0-12-381479-1

Slides are [here](#) (Links to an external site.).

Grading Policy

1. Exams

There will be two in-class closed notes, closed books tests (each with two hours). One sheet of review notes is allowed.

2. Homework

Five homework assignments are due before class of the scheduled date. *No late homework is accepted.* A submission after the due date receives zero point on that assignment. Each student should complete all the assignments *independently*.

3. Term Paper/Project

Each graduate student will conduct an investigation on given topics by writing a term paper and implementing some of data mining algorithms. More details will appear later.

4. Grading Scheme

Test I: 25%

Test II: 30%

Homework: 40%, 5 assignments each 8%

Attendance and Active Participation: 5% (anyone who misses 3 or more classes will lose this 5%)

For graduate students:

Graduate students who register in this course must write an additional term paper investigating a research topic in data mining algorithms or systems. The weight for this work is 20%, bringing the total for graduate students to 120%, which will be prorated to 100% to determine the grade. List of topics, resources, and guidelines for term papers will be provided in class.

The letter grade you will receive depends on the numerical weighted score and your overall performance in this course. Roughly,

90-100	A
80-89	B
70-79	C
60-69	D
< 60	F

5. Attendance Policy

Attendance is required for all the class meetings and tests. If you will be absent for an exam due to circumstances beyond control, let me know in advance so that we can arrange a make-up test. Students who miss an exam without informing me *prior to* the exam will receive a 0 on that exam.

6. Academic Honesty

The instructor will deal strictly with any violations of academic honesty and integrity in this course. See [Academic Integrity Policy \(Links to an external site.\)](#) for more details. Any student who violates this policy receives zero point for the work and is

subject to a reduction of the final grade of this course (up to "F"). The instructor will report the case to the university.

Please read the textbook sections in advance, attend all the lectures and start on homework early. If you have any difficulties, please see me during my office hours, email me, or make an appointment. I am more than happy to help you.

Lixin Fu