

CS 4710/6710 Introduction to Data Mining

Spring, 2025

Description: Introduction of basic concepts and methods of data mining (aka knowledge discovery from data).

Prerequisite: CS 3100 Data Structures and Algorithms, or equivalent.

Instructor: Dr. Soon M. Chung
403 Russ, 775-5119
soon.chung@wright.edu
<http://web1.cs.wright.edu/~schung>

Class: M. W. F. 10:10 –11:05 am, 146 Russ Center

Office hour: M. W. 2:40 am – 3:40 pm at 403 Russ or by appointment.

*use e-mail for short questions.

Text Book: Jiawei Han, Jan Pei, and Hanghang Tong, *Data Mining Concepts and Techniques*, 4th edition, 2023, Morgan Kaufmann (or Jiawei Han and Micheline Kamber, *Data Mining Concepts and Techniques*, 3rd edition, 2012, Morgan Kaufmann).

Course Topics:

Introduction (chapter 1)
Data types, statistics, distance and similarity measurements (chapter 2)
Pattern Mining: basic concepts and methods (Frequent itemsets, Association rules, and Correlations) (chapter 4) (chap. 6 in 3rd ed.)
Cluster Analysis: basic concepts and methods (chapter 8) (chap. 10 in 3rd ed.)
Data Preprocessing (chapter 2) (chap. 3 in 3rd ed.)
Classification: basic concepts (chapter 6) (chap. 8 in 3rd ed.)
Data Warehousing and Online Analytical Processing (OLAP) (chapter3) (chap. 4 in 3rd ed.)
Data Cube Technology (chapter3) (chap. 5 in 3rd ed.)

Reference Book: Charu C. Aggarwal, *Data Mining: The Textbook*, Springer, 2015

Grading: A:[85,100], B:[75,85), C:[65,75), D:[55,65), F:[0,55)

Midterm 30% (3/14, F. 10:10 –11:05 am), Final 35% (4/30, W. 10:15 am -12:15 pm), and Project 35%

Project: Implementation of a data mining algorithm 35%

{project plan 5%, description of the mining algorithm and dataset used 5%, implementation design and its description 5%, implementation and documentation 10%, in-class presentation 5%, discussion 5% }

- **Project plan is due on 2/26, Wednesday.**
- **Final report is due on 4/23, Wednesday.**