CS 428 Data Mining and Warehousing Fundamentals

Summer 2024

Course Credits: 4
Contact Hours: 56 hours
Instructor: TBA
Email: TBA

COURSE OBJECTIVES

This course provides a comprehensive overview of key concepts in data mining and data warehousing, covering fundamental aspects such as objectives, architectures, algorithms, implementations, and real-world applications. Students will gain insights into the pivotal role of data mining and data warehousing in extracting valuable knowledge from large datasets to support informed decision-making processes.

Upon Completion of this Course, students will be able to:

1. Understand the foundational principles of collecting, storing, and processing operational data to support day-to-day business activities;

2. Explore how data can be analyzed and processed to facilitate strategic decision-making within organizations;

3. Delve into the design, implementation, and utilization of data warehouses, as well as the role of OLAP in providing multidimensional analysis capabilities;

4. Explore clustering algorithms to group similar data points together, aiding in the identification of inherent structures within datasets;

5. Investigate techniques for identifying patterns and trends within datasets to characterize and describe the underlying data.

PREREQUISITES

CS 258 Data Structures and Algorithms; CS 300 Database System

GRADING

Grades will be determined by accumulating points, with 100 points being the

maximum, as follows:

ITEM	POINTS
2 Quizzes	20 Points
2 Exercises	20 Points
Midterm	30 Points
Final Project and Exam	30 Points
Total	100 Points

Late submissions will be graded at the end of the course. Grades will be assigned according to the following rule:

A ≥ 90 > B ≥ 80 > C ≥ 70 > D ≥ 60 > F.

We reserve the right to make adjustments to the overall grading policy.

COURSE MATERIALS

Required Texts:

Paulraj Ponniah, *Data Warehousing Fundamentals for IT Professionals*, 2nd Edition, Wiley-Blackwell, 2010.

Recommended (Optional) Texts or Other Materials:

None

COURSE TOPICS

MODULE	TASKS
Module 1	**Topics:** Topic 1: The Compelling Need For Data Warehousing Topic 2: Data Warehousing—The Only Viable Solution Topic 3: Architectural Types Topic 4: Centralized Data Warehouse **Assessments:** Quiz#1

Module 2	**Topics:** Topic 5: Genetic Algorithms Topic 6: Implementation Options Topic 7: Examples of Typical Implementations Topic 8: Data Warehouse Versus Operational Systems **Assessments:** Exercise#1
Module 3	**Topics:** Topic 9: Operational Infrastructure Topic 10: Processing Requirements in the New Environment Topic 11: Strategic Information from the Data Warehouse Topic 12: Real-Time Data Warehousing **Assessments:** Midterm Quiz#2
Module 4	**Topics:** Topic 13: Data Mining Basics Topic 14: Online Analytical Processing (OLAP) Topic 15: OLAP Versus Data Mining Topic 16: OLAP Definitions and Rules **Assessments:** Exercise#2
Module 5	**Topics:** Topic 17: Data Clustering Topic 18: Distinguishing Characteristics Topic 19: Applications in CRM Topic 20: Applications in the Telecommunications Industry **Assessments:** Final Exam Final Project

ATTENDANCE

1) Class attendance is required. Missing classes without permission will lead to decrease in overall grade.

Missing less than two classes: no penalty.

Missing more than two classes: 7% will be taken off from the overall grade.

If the instructor reports a student's frequent missing of class to the Soochow University Academic Administration Office, the student might get a written warning and might be prohibited from attending final exam.

2) Participants in this course are expected to arrive in class promptly and adequately prepared. The primary objective of this course is to critically engage with the readings and the subject matter. Therefore, course participants are expected to have completed the reading prior to class and prepare thoughtful reflections/commentaries to share with fellow colleagues.

LEARNING REQUIREMENTS

1) Late assignments are not acceptable and are subjected to grade deductions.

2) Assignments submitted in the wrong format will be counted as not submitted.

3) Failure to submit or fulfill any required course component results in failure of the class.

4) Make-up for midterm and final exams only with valid excuses, as defined by the University.

5) In order to earn a Certificate of Completion, participants must thoughtfully complete all assignments by stated deadlines and earn an average quiz score of 50% or greater.

TECHNOLOGY POLICY

The use of electronic devices in class is distracting, both for the user and for the rest of the class. Only non-programmable calculators can be used in the tests and exam. Any attempts to use cell phones and other electronic communication devices will be seemed as cheating. Laptops are discouraged, unless you use them for activities DIRECTLY related to the course (eg., note taking, reading course documents).

ACEDEMIC INTEGRITY POLICY

Soochow University highly values the academic integrity and aims to promote the academic fairness, honesty and responsibility. Any academic dishonesty behaviors and any attempts to cheats and plagiarism will be reported to the university administration office. A written warning and the relevant penalties will be imposed. The record might be shown on the official university transcript.

DISABILITY ACCOMMODATION

Soochow University is committed to maintaining a barrier-free environment so that students with disabilities can fully access programs, courses, services, and activities at Soochow University. Students with disabilities who require accommodations for access to and/or participation in this course are welcome.

Note:

Please contact the University Administrative Office immediately if you have a learning disability, a medical issue, or any other type of problem that prevents professors from seeing you have learned the course material.