

CS699OL Syllabus (Generic)

Note: This is a simplified, generic syllabus. Actual syllabus of each term is slightly different from this one but overall content is the same (unless there is a significant course update).

Course Description

The goal of this course is to study basic concepts and techniques of data mining. The topics include data preparation, classification, performance evaluation, association rule mining, and clustering. We will discuss basic data mining algorithms in the class and students will practice data mining techniques using data mining software. Students will use Weka and Oracle or JMP Pro.

Course Overview

- Introduction to data mining
- Data Preprocessing
- Classification
- Association Rule Mining
- Clustering
- Case Studies

Module 1 – Introduction and Software Installation

- Overview of Data Mining and Data Warehousing
- Installation of Weka, Oracle/JMP Pro (data mining tools)

Module 2 – Data Exploration and Preprocessing

- Data Exploration
- Data Preparation

Module 3 – Classification

- Naïve Bayes
- Decision Tree

Module 4 – Performance Evaluation and Other Classifiers

- Performance Evaluation
- Other Classifiers

Module 5 – Association Analysis

- Association Analysis
- Association Rule Evaluation

Module 6 – Clustering and Case Studies

- Clustering
- Case Studies

Required Book

Han, J. & Kamber, M. (2012). *Data Mining: Concepts and Techniques* (3rd ed.). Morgan Kaufmann Publishers.

Recommended Book

- Ian H. Witten, E. Frank, and M.A. Hall, "Data Mining Practical Machine Learning Tools and Techniques," Fourth Ed., 2016, Morgan Kaufmann.
- G. Shmueli, P.C. Bruce, M.L. Stephens, N.R. Patel, 'Data Mining for Business Analytics; Concepts, Techniques, and Applications with JMP PRO,' Wiley, 2017

Supplemental Material

Supplemental Material is available on Blackboard. Download the zip file containing the supplementary materials to your local hard disk.

Study Guide

Module 1 Study Guide and Deliverables

Readings: *Online Lectures*

Data Mining: Concepts and Techniques: Chapter 1, Section 4.1

Discussions: Discussion 1

Assignments: Assignment 1

Assessments: Quiz 1

Module 2 Study Guide and Deliverables

Readings: Online Lectures

Data Mining: Concepts and Techniques:

Chapter 2 –Getting to Know Your Data: Sections 2.1, 2.2, 2.3, 2.4.1, 2.4.2, 2.4.3, 2.4.4, 2.4.5, 2.4.7

Chapter 3 – Data Preprocessing: Sections 3.1, 3.2, 3.3, 3.4.1, 3.4.6, 3.4.7, 3.4.8, 3.4.9, 3.5

Discussions: Discussion 2

Assignments: Assignment 2

Assessments: Quiz 2

Module 3 Study Guide and Deliverables

Readings: Online Lectures

Data Mining: Concepts and Techniques

Chapter 8 – Classification: Basic Concepts: Sections 8.1, 8.2.1, 8.2.2, 8.2.3, 8.3, 8.4.1, 8.4.2

Discussions: Discussion 3

Assignments: Assignment 3

Assessments: Quiz 3

Module 4 Study Guide and Deliverables

Readings: Online Lectures

Data Mining: Concepts and Techniques

Sections 8.5.1, 8.5.2, 8.5.3, 8.5.4, 8.5.5, 8.6.1, 8.6.2, 9.5.1

Discussions: Discussion 4

Assignments: Assignment 4, Project assigned.

Assessments: Quiz 4

Module 5 Study Guide and Deliverables

Readings: Online Lectures

Data Mining: Concepts and Techniques

Chapter 6 – Association Analysis: Sections 6.1, 6.2.1, 6.2.2, 6.3.1, 6.3.2, 6.3.3

Discussions: Discussion 5

Assignments: Assignment 5

Assessments: Quiz 5

Module 6 Study Guide and Deliverables

Readings: Online Lectures

Data Mining: Concepts and Techniques

Chapter 10 Cluster Analysis - Sections 10.1.1, 10.1.2, 10.1.3, 10.2.1, 10.3.1, 10.3.2, 10.6

Discussions: Discussion 6

Assignments: Assignment 6, Project due.

Final Exam Details

The exam is a three-hour open-book exam consisting of 20 multiple-choice questions. All 20 questions are calculation questions and you need to bring a calculator with a "log" function.

Grading Structure and Distribution

The grade for the course is determined by the following:

Overall Grading Percentages

Assignments 30%

Discussions 10%

Project 10%
Quizzes 20%
Proctored Final Examination 30%

$94 \leq G$ A
 $90 \leq G < 94$ A-
 $87 \leq G < 90$ B+
 $83 \leq G < 87$ B
 $80 \leq G < 83$ B-
 $77 \leq G < 80$ C+
 $73 \leq G < 77$ C
 $70 \leq G < 73$ C-
 $60 \leq G < 70$ D
 $G < 60$ F

Live Classroom Discussion and Archive

The instructor will hold six synchronous Live Classrooms in Weeks 1 – 6, and they will be archived for further viewing. Your participation, while not mandatory, will be valuable to you and the class.

Discussions

Graded Discussions—Students will be participating in discussions that will be graded on a 100-point scale.

Quizzes

There will be five quizzes, one per every week except in Week 6 and Week 7. The primary goal of quizzes is to let students keep current with the course material.

Proctored Final Exam

There will be a proctored Final Exam for this course. You will be responsible for scheduling your own appointment with an approved proctoring option. Detailed instructions about setting up an appointment will be forthcoming from the proctored exam coordinator.

Delays

If, for any reason, you are unable to meet any deadline, contact your Course Facilitator. Assignments are expected to be submitted by their respective due dates. Extensions may be granted, though **only under mitigating circumstances**. If your facilitator grants an extension, you will not be penalized. If you submit an assignment late without a permission of your facilitator, there will be a late penalty of 10% per day.

Quiz Details

- There are 20 questions per quiz. You can access the quiz details from the assessments menu.
- The questions are either multiple choice or True/False.
- All questions are randomized.
- The points for each question are shown.
- The quiz questions will display one at a time on your screen.
- You may skip over questions and revisit them in any order.
- You will have 90 minutes to complete the quiz. Also note:
 - You can take each quiz only once.
 - You may not pause the quiz and return to it later.
 - You will be able to continue to save answers to questions after the time has expired, but any late answers will be time stamped and marked as late. This will allow us to grade your quiz fairly in the event that technical difficulties occur while you take your quiz.

Comments on the Quiz

There will be a short-answer area at the end of the quiz; it appears as a quiz question, but there are no points for this item. Use this as a place to provide feedback about the quiz as a whole or to comment upon a particular quiz item. Be sure to reference the question number. Your facilitator will examine your comments in order to decide whether a grade adjustment or other action should be taken. If a technical issue of any kind arises during the exam, complete the exam, answering the remaining questions, and then contact your facilitator or instructor immediately.