

DS 479: Data Mining and Machine Learning

School of Technology & Computing

5 Credits, Undergraduate Course

Summer 2022

Grading Type: Decimal

Access to the Internet is required.

All written assignments must be in Microsoft-Word-compatible formats.

See the library's APA Style Guide tutorial for a list of resources that can help you use APA style.

Faculty Information

Professional experience information for instructors is found under *Faculty Information* in the online course menu.

Contact Information

Contact information for instructors is found under *Faculty Information* in the online course menu.

Email: [first name] [last name]

Phone: [xxx-xxx-xxxx]

Office Hours and Response Time: [I am available through MS Teams xxday and xxday nights between xx p.m.– xx p.m. I will respond within 24 hours. I will grade within 3 business days after the due date.]

Bio: (keep images under 300px wide)

Course Description

Data mining involves the processing, analysis, and presentation of data to gain valuable information. Machine learning refers to a broad set of algorithms for identifying patterns in data to build models that might then be possibly productized. Students completing this course will develop an understanding of data mining concepts such as proximity measurement, data preparation, cluster analysis, classification and regression and apply machine learning algorithms such as supervised learning, unsupervised learning, and deep learning.

Course Resources

Required and recommended resources to complete coursework and assignments are found on the course Reading List. Note: resources listed under "Required - Must Purchase" should be purchased from a vendor of the student's own choosing; resources listed under "Available from the Library" are available at no cost to students.

Course Outcomes

As a result of this course, students will know or be able to do the following:

- Understand the bias-variance tradeoff in supervised learning.
- Understand the importance of feature selection for clustering, classification, and regression.
- Apply clustering, classification, and regression algorithms to small and medium data sets.
- Analyze the performance of supervised and unsupervised algorithms using various metrics.

- Evaluate criteria that might lead to selection of one method over another in supervised learning.
- Create a data mining and machine learning deliverable using appropriate algorithms.

Grading Scale

The grades earned for the course will be calculated using City University of Seattle’s decimal grading system, found in the current University Catalog (<https://www.cityu.edu/catalog/>).

Grading rubrics with details on how each assignment will be graded are located under *Assignments* and/or in *My Grades* in the online course menu. Students should review the rubric for each assignment prior to completing their work in order to understand how it will be assessed.

OVERVIEW OF REQUIRED ASSIGNMENTS	% OF FINAL GRADE	POINTS
The Muddiest Point (MP)	5%	50 = 5 points* 10 modules
Concept Test (CT)	5%	50 = 5 points* 10 modules
Discussion Board (DB)	10%	100 = 10 points* 10 modules
Hands-On Practice (HOP)	20%	200= 20 points* 10 modules
Programming Exercise (PE)	30%	300= 30 points * 10 modules
Knowledge Check (KC)	10%	100= 10 points * 10 modules
Team Project (TP)	20%	Proposal: 30 points Progress: 70 points Final Report: 70 points Final PPT: 30 points Subtotal: 200 points
TOTAL	100%	1,000 points

Course Assignments and Grading

The instructor will provide grading rubrics that will provide more detail as to how this assignment will be graded

The Muddiest Point (MP)

Before class, students are required to submit the Muddiest Point (MP) activity. The purpose of this activity is to stimulate student engagement. The instructor uses the MP to assess how students understood the required readings. The instructor also uses the MP to customize the lecture scope to implement Just-in-Time Teaching (JiTT). The MP consists of writing a brief reflective essay (<= 50 words) identifying the most confusing part (i.e., the MP) of the content covered in the upcoming module. If a student understood all concepts, the student needs to explain the most exciting aspect. There is one multiple-choice question from the required reading to demonstrate that the student understood the required readings.

Criteria	% of Grade
Participation	40%

Writing	40%
Accuracy	20%
TOTAL	100%

Concept Test (CT)

The instructor poses a problem based on key concepts of a lecture. After reflecting on the problem, students submit their response and the instructor review them without providing a correct answer. Students discuss their thought process and solution with a peer. Students then commit to an answer and re-submits their responses. Instructor reviews responses and thought processes with the correct answer.

Criteria	% of Grade
Engagement	100%
TOTAL	100%

Discussion Board (DB)

A student posts an answer to a weekly discussion topic in Discussion Board. The student also posts a response to two other students' posts by the end of each module. Comments and questions should be clear and thoughtful, with correct grammar, spelling, and punctuation. The instructor will grade the quality of your discussion postings on both content and response.

Components	% of Grade
Participation	50%
Writing	50%
TOTAL	100%

Hands-on Practice (HOP)

The instructor will assign Hands-on Practice exercises to a pair of students in class or individually online. Students pair up and practice exercises to learn specific programming languages, application programming interfaces (APIs), or tools related to the programming assignments or virtual labs.

Criteria	% of Grade
Practice Exercise	80%
Engagement	20%
TOTAL	100%

Programming Exercise (PE)

The students must individually perform the programming exercise, which is based on the topics and Hands-on Practice. No code sharing or copying from other sources are allowed. Non-executable programs will not be graded. The programs in poor coding styles will be asked to be resubmitted. **Please note that copying a segment of code from the Internet and submitting it as your work is considered as plagiarism.**

Criteria	% of Grade
Program Execution	40%
User Requirement	40%

Program Documentation	20%
TOTAL	100%

Knowledge Check (KC)

Weekly quizzes measure knowledge concepts acquired. Focus on the underlying principles and concepts rather than memorization to solve the quizzes.

Criteria	% of Grade
Correctness	100%
TOTAL	100%

Team Project (TP)

Teams consist of three to four students. Each team will use an instructor-approved topic relevant to the course.

The paper is to be between 6 and 7 pages. The required template for class submissions comes from international organizations, the Education Special Interest Group and the Computing Education + Information systems Applied Research. ([EDSIG/CONISAR](#)). The instructor may recommend teams submit their paper to conferences. Submissions are optional and will not impact the course grade. Additional revisions may be required after the course.

Three report templates and one presentation template are provided. The file name consists of team project number, team number, and the list of your team members. For example, "TP01 T03 Sam John Mark."

- TP01 for the proposal - "TP01 T0X Author1 Author2 Author3.docx"
- TP02 for the progress report - "TP02 T0X Author1 Author2 Author3.docx"
- TP03 for the final report - "TP03 T0X Author1 Author2 Author3.docx"
- TP04 for the final presentation slide - "TP04 T0X Author1 Author2 Author3.pptx"

As in any scholarly writing, students should not merely copy information from another author. Students should use evidence to support the contentions they have drawn from their findings and critically analyze related literature. In essence, each paper needs to be an analytical paper, not a summary of readings.

In addition, a team slide deck presentation is required.

- The presentation consists of 15+4 slides: 15 slides for content and 4 slides for cover, agenda, key reference, and Q&A.
- The PPT template is provided. Your team can change design and color.
- A presentation video (15 minutes) is required.
- A demo video (a maximum of 1-2 minutes) may be included. The demo time is included in the 15 minutes presentation.

Four submissions towards finalizing the Team Project are required according to the following schedule:

- Proposal (1 page; 30 points) - Starting (Module 1) & Ending (Module 3)
- Progress Report (3-4 pages; 70 points; graded after the proposal has been submitted) - Starting (Module 4) & Ending (Module 7)

- Final Report (6-7 pages; 70 points; graded after the progress has been submitted) - Starting (Module 8) & Ending (Module 10)
- Final PPT (15+4slides, 30 points; graded after the final report has been submitted) - Starting (Module 8) & Ending (Module 10)

Students are expected to use the assigned readings, videos, and other materials throughout the quarter. Students will need to utilize additional sources that were not assigned by the professor. While stylized after an industry report, nonetheless, students are expected to employ APA formatting of citations, footnotes, and bibliography. Students must cite the sources of all ideas, facts, and information used that are not their own, even if they have put the information into their own words. Failure to properly cite the source is plagiarism, even if the oversight is unintentional. To align with university academic integrity standards, check <https://library.cityu.edu/howto/apa-writing/avoid-plagiarism/>.

Project Description: Data Mining and Machine Learning Algorithms

Students are required to demonstrate the mastery of data mining and machine learning algorithms using a scripting language determined by the instructor.

Team Project (TP) Report

The student will provide a report formatted based on a template provided by the instructor. Students are required to improve the writing iteratively and incrementally every week. The revision will always happen during a quarter. Students will add new required sections to the existing paper every week.

The final report is the culmination of applied research and activities conducted throughout the quarter. The final report/paper provides a detailed problem and its solution likely to be encountered by a company or organization described in a case study supplied by the student.

Criteria	% of Grade
Structure	20%
Content	30%
Writing	30%
Reference	10%
Collaboration	10%
TOTAL	100%

Team Project (TP) Presentation

The student will report on the research outcomes, development, or other project efforts to an academically appropriate committee in a public forum. The nature of the presentation content will determine the specific makeup of the audience. The student will choose the format of the presentation, in consultation with the advisor. The layout and design must be appropriate and adequate to represent the outcomes of the effort. While students must make some form of a visual presentation, the presentation of the results may include publishing in a refereed publication, publication in a trade or popular magazine or journal, broadcast in an appropriate medium, or, in exceptional cases, limited dissemination within a closed community.

Each presenter will have 15 minutes for presentation and 5 minutes for questions and answers. Each presenter must keep the total presentation time limit strictly.

Criteria	% of Grade
Structure	20%

Visual Presentation	30%
Verbal Quality & Engagement	30%
Team Collaboration	20%
TOTAL	100%

COURSE POLICIES

Course policies on topics such as *Late Assignments*, *Participation*, and *Professional Writing* are found under *Course Information* in the online course menu. Students are responsible for reviewing and applying these policies while enrolled in this course.

University Policies

Students are responsible for understanding and adhering to all of City University of Seattle’s academic policies. The most current versions of these policies can be found in the [University Catalog](#) that is linked from the CityU Web site.

Title IX Statement

City University of Seattle and its faculty are committed to supporting our students and seeking an environment that is free of bias, discrimination, and harassment. If students have encountered any form of sexual misconduct (e.g. sexual assault, sexual harassment, stalking, domestic or dating violence), we encourage them to report this to the University. If a student speaks with a faculty member about an incident of misconduct, that faculty member must notify CityU’s Title IX coordinator and share the basic fact of the experience. The Title IX coordinator will then be available to assist students in understanding all of the options and in connecting students with all possible resources on and off campus.

To view CityU’s sexual misconduct policy and for resources, please visit the [Title IX](#) and [Campus Safety](#) pages in the my.cityu.edu portal.

Religious Accommodations

Washington state law requires that City University of Seattle develop a policy for accommodation of student absences or significant hardship due to reasons of faith or conscience, or for organized religious activities. The University’s policy, including more information about how to request an accommodation, is available in the University Catalog. Accommodations must be requested within the first two weeks of this course using the Religious Accommodations Request Form found on the student dashboard in the my.cityu.edu student portal.

Academic Integrity

Academic integrity in students requires the pursuit of scholarly activity that is free from fraud, deception and unauthorized collaboration with other individuals. Students are responsible for understanding CityU’s policy on academic integrity and adhering to its standards in meeting all

course requirements. A complete copy of this policy can be found in the [University Catalog](#) under *Student Rights and Responsibilities* on the page titled *Academic Integrity Policy*.

Attendance

Students taking courses in any format at the University are expected to be diligent in their studies and to attend class regularly.

Regular class attendance is important in achieving learning outcomes in the course and may be a valid consideration in determining the final grade. For classes where a physical presence is required, a student has attended if they are present at any time during the class session. For online classes, a student has attended if they have posted or submitted an assignment. A complete copy of this policy can be in the [University Catalog](#) under *Student Rights and Responsibilities* on the page titled *Attendance*.

Support Services

Disability Services Accommodations Statement

Students with a documented disability who wish to request academic accommodations are encouraged to contact Disability Support Services to discuss accommodation requests and eligibility requirements. Please contact Disability Support Services at disability@cityu.edu or 206.239.4752 or visit the [Disability Support Services](#) page in the my.cityu.edu portal.

Confidentiality will be observed in all inquiries. Once approved, information about academic accommodations will be shared with course instructors.

Library Services

CityU librarians are available to help students find the resources and information they need to succeed in this course. Contact a CityU librarian through the [Ask a Librarian](#) service, or access [library resources and services online](#), 24 hours a day, seven days a week.

Smarthinking Tutoring

CityU students have access to free online tutoring offered through Smarthinking, including writing support, from certified tutors 24 hours a day, seven days a week. Contact CityU's Student Support Center at help@cityu.ed to request a user name and password.