

CSE 5243-0050 – Spring 2021 Introduction to Data Mining

Time/Place: 12:45 PM-02:05 PM, Zoom Meetings

Instructor: Dr. Ping Zhang; 310A LT; zhang.10631@osu.edu; <http://pingzhang.net>

Teaching Assistant (TA): Caroline Shin; shin.768@osu.edu

Level and credits: U/G, 3

Office Hours: (Instructor) Wed 11:00 AM – 12:00 PM; (TA) Mon 1:00 PM – 2:00 PM

Course Description: Knowledge discovery, data mining, data preprocessing, data transformations; clustering, classification, frequent pattern mining, anomaly detection, avoiding false discoveries, graph and network analysis; applications.

Prerequisites: Introduction to Databases, Introduction to Algorithms, or grad standing or permission of instructor

Textbooks:

- (Required) Jiawei Han, Micheline Kamber and Jian Pei, Data Mining: Concepts and Techniques (3rd ed), 2011
- (Required) Pang-Ning Tan, Michael Steinbach, and Vipin Kumar, Introduction to data Mining (1st or 2nd ed), 2006
- (Optional) Charu C. Aggarwal, Data Mining: The Textbook, Springer, 2015.
- (Optional) Mohammed J. Zaki and Wagner Meira Jr., Data Mining and Analysis: Fundamental Concepts and Algorithms, 2014.

Grading Plan:

- Quizzes: 10%
 - Graded 10 quizzes from 12 opportunities
- Homework: 30%
 - 3 programming assignments
- Data Science Challenge: 5%
 - Winning Team Bonus: up to 1.5%
- Group Presentation: 5%
- Midterm exam: 25% (Friday March 5 in class)
- Final exam: 25% (TBD)

Health and safety requirements

All students, faculty and staff are required to comply with and stay up to date on all university safety and health guidance (<https://safeandhealthy.osu.edu>), which includes wearing a face mask in any indoor space and maintaining a safe physical distance at all times. Non-compliance will be warned first and disciplinary actions will be taken for repeated offenses.

Carmen access

You will need to use [BuckeyePass](#) multi-factor authentication to access your courses in Carmen. To ensure that you are able to connect to Carmen at all times, it is recommended that you take the following steps:

- Register multiple devices in case something happens to your primary device. Visit the [BuckeyePass - Adding a Device](#) help article for step-by-step instructions.
- Request passcodes to keep as a backup authentication option. When you see the Duo login screen on your computer, click **Enter a Passcode** and then click the **Text me new codes** button that appears. This will text you ten passcodes good for 365 days that can each be used once.
- Download the [Duo Mobile application](#) to all of your registered devices for the ability to generate one-time codes in the event that you lose cell, data, or Wi-Fi service.
- If none of these options will meet the needs of your situation, you can contact the IT Service Desk at 614-688-4357 (HELP) and IT support staff will work out a solution with you.

Course Technology Resources

For help with your password, university email, Carmen, or any other technology issues, questions, or requests, contact the Ohio State IT Service Desk. Standard support hours are available at ocio.osu.edu/help/hours, and support for urgent issues is available 24/7.

- **Self-Service and Chat support:** ocio.osu.edu/help
- **Phone:** 614-688-4357(HELP)
- **Email:** servicedesk@osu.edu
- **TDD:** 614-688-8743

Baseline technical skills for online courses

- Basic computer and web-browsing skills
- Navigating Carmen: for questions about specific functionality, see the [Canvas Student Guide](#).

Required Technology skills specific to this course

- [CarmenZoom virtual meetings](#)

Required equipment

- Computer: current Mac (OS X) or PC (Windows 7+) with high-speed internet connection
- Webcam: built-in or external webcam, fully installed and tested
- Microphone: built-in laptop or tablet mic or external microphone
- Other: a mobile device (smartphone or tablet) or landline to use for BuckeyePass authentication

Required software

- [Microsoft Office 365](#): All Ohio State students are now eligible for free Microsoft Office 365 ProPlus through Microsoft's Student Advantage program. Full instructions for downloading and installation can be found [at go.osu.edu/office365help](http://go.osu.edu/office365help).

Copyright disclaimer

The materials used in connection with this course may be subject to copyright protection and are only for the use of students officially enrolled in the course for the educational purposes associated with the course. Copyright law must be considered before copying, retaining, or disseminating materials outside of the course.

Ohio State Academic Integrity Policy

Academic integrity is essential to maintaining an environment that fosters excellence in teaching, research, and other educational and scholarly activities. Thus, The Ohio State University and the Committee on Academic Misconduct (COAM) expect that all students have read and understand the university's [Code of Student Conduct](#), and that all students will complete all academic and scholarly assignments with fairness and honesty. Students must recognize that failure to follow the rules and

guidelines established in the university's *Code of Student Conduct* and this syllabus may constitute "Academic Misconduct."

The Ohio State University's *Code of Student Conduct* (Section 3335-23-04) defines academic misconduct as: "Any activity that tends to compromise the academic integrity of the university or subvert the educational process." Examples of academic misconduct include (but are not limited to) plagiarism, collusion (unauthorized collaboration), copying the work of another student, and possession of unauthorized materials during an examination. Ignorance of the university's *Code of Student Conduct* is never considered an excuse for academic misconduct, so I recommend that you review the *Code of Student Conduct* and, specifically, the sections dealing with academic misconduct.

If I suspect that a student has committed academic misconduct in this course, I am obligated by university rules to report my suspicions to the Committee on Academic Misconduct. If COAM determines that you have violated the university's *Code of Student Conduct* (i.e., committed academic misconduct), the sanctions for the misconduct could include a failing grade in this course and suspension or dismissal from the university.

If you have any questions about the above policy or what constitutes academic misconduct in this course, please contact me.

Other sources of information on academic misconduct (integrity) to which you can refer include:

- The Committee on Academic Misconduct web pages ([COAM Home](#))
- *Ten Suggestions for Preserving Academic Integrity* ([Ten Suggestions](#))
- *Eight Cardinal Rules of Academic Integrity* (www.northwestern.edu/uacc/8cards.htm)

Statement on Title IX

All students and employees at Ohio State have the right to work and learn in an environment free from harassment and discrimination based on sex or gender, and the university can arrange interim measures, provide support resources, and explain investigation options, including referral to confidential resources.

If you or someone you know has been harassed or discriminated against based on your sex or gender, including sexual harassment, sexual assault, relationship violence, stalking, or sexual exploitation, you may find information about your rights and options at titleix.osu.edu or by contacting the Ohio State Title IX Coordinator at titleix@osu.edu. Title IX is part of the Office of Institutional Equity (OIE) at Ohio State, which responds to all bias-motivated incidents of harassment and discrimination, such as race, religion, national origin and disability. For more information on OIE, visit equity.osu.edu or email equity@osu.edu.

Your Mental Health

As a student you may experience a range of issues that can cause barriers to learning, such as strained relationships, increased anxiety, alcohol/drug problems, feeling down, difficulty concentrating and/or lack of motivation. These mental health concerns or stressful events may lead to diminished academic performance or reduce a student's ability to participate in daily activities. The Ohio State University offers services to assist you with addressing these and other concerns you may be experiencing. If you find yourself feeling isolated, anxious or overwhelmed, please know that there are resources to help: ccs.osu.edu. You can reach an on-call counselor when CCS is closed at (614) 292-5766 and 24 hour emergency help is also available through the 24/7 National Prevention Hotline at 1-(800)-273-TALK or at suicidepreventionlifeline.org. The Ohio State Wellness app is also a great resource available at go.osu.edu/wellnessapp.

Course Syllabus:

Week	Date	Topic	Assignment	Quiz
1	01/13/2021	Introduction		
1	01/15/2021	Introduction		
2	01/20/2021	Data & Data Preprocessing		Quiz 1
2	01/22/2021	Data & Data Preprocessing		
3	01/27/2021	Classification:Basic Concepts/Methods		Quiz 2
3	01/29/2021	Classification:Basic Concepts/Methods	Assignment 1 Out	
4	02/03/2021	Classification:Basic Concepts/Methods		Quiz 3
4	02/05/2021	Classification:Basic Concepts/Methods		
5	02/10/2021	Classification:Advanced Methods		Quiz 4
5	02/12/2021	Classification:Advanced Methods	Competition Out	
6	02/17/2021	Association Analysis		Quiz 5
6	02/19/2021	Association Analysis	Assignment 2 Out	
7	02/24/2021	No Class - Instructional Break		
7	02/26/2021	Association Analysis		Quiz 6
8	03/03/2021	Midterm Preparation	Assignment 1 Due	
8	03/05/2021	Midterm Exam (In Class)		
9	03/10/2021	Association Analysis + Midterm Rev		
9	03/12/2021	Association Analysis: Advanced Concepts		Quiz 7
10	03/17/2021	Association Analysis: Advanced Concepts		
10	03/19/2021	Cluster Analysis: Basic Concepts and Algorithms	Assignment 3 Out	Quiz 8
11	03/24/2021	Cluster Analysis: Basic Concepts and Algorithms	Assignment 2 Due	
11	03/26/2021	Cluster Analysis: Basic Concepts and Algorithms		Quiz 9
12	03/31/2021	No Class - Instructional Break		
12	04/02/2021	Cluster Analysis: Additional Issues and Algorithms		
13	04/07/2021	Cluster Analysis + Anomaly Detection		Quiz 10

13	04/09/2021	Group Presentation: Classification		
14	04/14/2021	Group Presentation: Association Analysis		Quiz 11
14	04/16/2021	Group Presentation: Clustering	Competition Due	
15	04/21/2021	Group Presentation: Anomaly Detection and Others	Assignment 3 Due	Quiz 12
15	04/23/2021	Data Competition + Final Preparation		
16	TBD	Final Exam		