



ISE 535: Data Mining
4 Units

Day/Time: Tues/Thurs 2:00 – 3:50PM, Spring 2025
Location: LVL 17

Instructor: Dr. Bruce Wilcox
Office: GER 203
Office Hours: Monday: 4:00PM – 4:45PM
Wednesday: 3:00PM – 4:45PM

Contact Info: brucewil@usc.edu

Teaching Assistants: Shijie Gao, Zuhayer Mahtab

Catalog Description

Data preprocessing, cleaning and visualization, statistical inference and testing, data classification, clustering analysis, association pattern mining, anomaly detection and false discoveries.

Course Description

Data mining is the discipline of extracting useful insights from large quantities of data. As such, the focus in this class is on statistical inference but not on prediction (which is the topic of ISE-529).

This course is organized into two broad sections:

- Exploratory data analysis and statistical data analysis techniques to find useful information from data.
- Algorithm-based data mining techniques, statistical (machine) learning-based techniques for data classification, clustering analysis, and association pattern mining.

To the maximum extent possible, this course teaches the concepts by means of case studies with real-world data and examples.

Learning Objectives and Outcomes

- Develop an advanced level of proficiency with the preprocessing, visualization, and statistical analysis of data as well as the primary data mining algorithmic techniques.
- Review and re-enforce basic statistical concepts that are important in the field of data science.
- Apply unsupervised modeling techniques (including clustering and association rule mining) to analyze and obtain insights from data.
- Take raw data and perform all the steps necessary to generate a professional exploratory data analysis report.

Class Delivery Mode: This class will be delivered via classroom lectures, homework assignments, exams, and a final project. Students are expected to attend lectures in person. Exams must be taken in person.

Prerequisite(s): None

Recommended Preparation:

Undergraduate course in statistics and working knowledge of Python.

Course Policies

All course materials (lecture slides, homework, exercises, etc.) will be distributed via Brightspace. They are only for the students who are enrolled in the course. Do **NOT** post or duplicate the course materials online or distribute them electronically or in print without prior explicit permission from the instructor. All assignments will be submitted through Gradescope.

Technological Proficiency and Hardware/Software Required

This course will utilize the Python programming language and associated libraries which are open source and available to the students for no cost.

Textbooks

The theoretical materials in the course are drawn from the following texts:

- Pang-Ning Tan, et. al., Introduction to Data Mining 2nd ed. 2019 ISBN 978-0-13-312890-1
- Charu C. Aggarwal, Data Mining, Springer, 2015 ISBN 9783319141411
- Bruce, et. al., Practical Statistics for Data Scientists, O'Reilly, 2020 (PSDS)

Grading Breakdown

The course grade distribution is as follows:

- Homework assignments (8-9): 30% of final grade
- Mid-terms (2): 40% of final grade
- Final project: 30% of final grade

The two mid-term exams will be held in-person during the class time. They will be closed book with one-page of cheat sheet permitted.

Grading Scale

Course final grades will be determined using the following scale:

A	95-100
A-	90-94.9
B+	87-89.9
B	83-86.9
B-	80-82.9
C+	75-79.9
C	70-74.9
C-	50-69.0
F	Below 50

Borderline averages between two letter grades may be rounded up based on class engagement at the instructor's discretion.

Assignment Submission Policy, Timelines, and Rules for Submission

- Assignments will be posted on Brightspace and submitted for grading on GradeScope (student instructions will be provided).
- Assignments turned in after the due date will be penalized 10%. Assignments not turned in within 48 hours of the due date will not be accepted.
- If you feel that your submission was not graded correctly, use the GradeScope "Request Regrade" button and type an explanation of why you feel the grade should be changed. I will accept requests for one week after grades published, after which the request regrade function will be disabled.
- The lowest homework grade for the semester will be dropped from the final grade computation.

Use of Generative AI in this Course

Generative AI is encouraged: You are expected to use AI (e.g., ChatGPT and image generation tools) in this class. Learning to use AI is an emerging skill; this is an opportunity for you to discuss with the instructor appropriate use of these tools. Keep in mind the following:

- AI tools are permitted to help you brainstorm topics or revise work you have already written.
- If you provide minimum-effort prompts, you will get low-quality results. You will need to refine your prompts to get good outcomes. This will take work.
- Proceed with caution when using AI tools and do not assume the information provided is accurate or trustworthy. If it gives you a number or fact: assume it is incorrect unless you either know the correct answer or can verify its accuracy with another source. You will be responsible for any errors or omissions provided by the tool. It works best for topics you understand.

Course Schedule

Week	W/E	Topics	Assignments
1	1/17	Module 1: Introduction to Data Mining	
2	1/24	Module 2: Exploratory Data Analysis	Module 2A HW Assigned
3	1/31		Module 2A HW Due Module 2B HW Assigned
4	2/7	Module 3: Statistical Data Analysis	Module 2B HW Due Module 3A HW Assigned
5	2/14		Module 3A HW Due Module 3B HW Assigned
6	2/21	Mid-term 1 (2/20)	Module 3B HW Due
7	2/28	Module 4: Inference and Interpretability in Supervised Learning	
8	3/14		Module 4 HW Assigned
9	3/21	Module 5: Unsupervised Learning Methods	Module 4 HW Due Module 5A HW Assigned
10	3/28		Module 5A HW Due Module 5B HW Assigned
11	4/4		Module 5B HW Due Module 5C HW Assigned
12	4/11	Module 6: Association Rule Mining	Module 5C HW Due
13	4/18		Module 6 HW Assigned
14	4/25	Mid-term 2 (4/22)	Module 6 HW Due Final project assigned
15	5/2	Class review / final project preparation	
		Final project due – 5/8	

Notes:

- The “W/E” column above means “week ending” and the date listed is the Friday of that week and the corresponding row refers to topics and assignments covered on one or both class days (Tuesdays or Thursdays) of that week.
- This schedule is subject to changes throughout the semester. This syllabus will not be updated, but the latest schedule will always be available on Piazza.
- The current official due dates for the homework assignments can be found in GradeScope.

Statement on Academic Conduct and Support Systems

Academic Conduct:

Plagiarism – presenting someone else’s ideas as your own, either verbatim or recast in your own words – is a serious academic offense with serious consequences. Please familiarize yourself with the discussion of plagiarism in SCampus in Part B, Section 11, “Behavior Violating University Standards” policy.usc.edu/scampus-part-b. Other forms of academic dishonesty are equally unacceptable. See additional information in SCampus and university policies on scientific misconduct, policy.usc.edu/scientific-misconduct.

Discrimination, sexual assault, and harassment are not tolerated by the university. You are encouraged to report any incidents to the Office of Equity and Diversity <http://equity.usc.edu> or to the Department of Public Safety <http://capsnet.usc.edu/department/department-public-safety/online-forms/contact-us>. This is important for the safety of the whole USC community. Another member of the university community – such as a friend, classmate, advisor, or faculty member – can help initiate the report, or can initiate the report on behalf of another person. The Center for Women and Men <http://www.usc.edu/student-affairs/cwm/> provides 24/7 confidential support, and the sexual assault resource center webpage <http://sarc.usc.edu> describes reporting options and other resources.

Support Systems:

Student Health Counseling Services - (213) 740-7711 – 24/7 on call
engemannshc.usc.edu/counseling

Free and confidential mental health treatment for students, including short-term psychotherapy, group counseling, stress fitness workshops, and crisis intervention.

National Suicide Prevention Lifeline - 1 (800) 273-8255 – 24/7 on call
suicidepreventionlifeline.org

Free and confidential emotional support to people in suicidal crisis or emotional distress 24 hours a day, 7 days a week.

Relationship and Sexual Violence Prevention Services (RSVP) - (213) 740-4900 – 24/7 on call
engemannshc.usc.edu/rsvp

Free and confidential therapy services, workshops, and training for situations related to gender-based harm.

Office of Equity and Diversity (OED) | Title IX - (213) 740-5086
equity.usc.edu, titleix.usc.edu

Information about how to get help or help a survivor of harassment or discrimination, rights of protected classes, reporting options, and additional resources for students, faculty, staff, visitors, and applicants. The university prohibits discrimination or harassment based on the following protected characteristics: race, color, national origin, ancestry, religion, sex, gender, gender identity, gender expression, sexual orientation, age, physical disability, medical condition, mental disability, marital status, pregnancy, veteran status, genetic information, and any other characteristic which may be specified in applicable laws and governmental regulations.

Bias Assessment Response and Support - (213) 740-2421
studentaffairs.usc.edu/bias-assessment-response-support

Avenue to report incidents of bias, hate crimes, and microaggressions for appropriate investigation and response.

The Office of Disability Services and Programs - (213) 740-0776

dsp.usc.edu

Support and accommodations for students with disabilities. Services include assistance in providing readers/notetakers/interpreters, special accommodations for test taking needs, assistance with architectural barriers, assistive technology, and support for individual needs.

USC Support and Advocacy - (213) 821-4710

studentaffairs.usc.edu/ssu

Assists students and families in resolving complex personal, financial, and academic issues adversely affecting their success as a student.

Diversity at USC - (213) 740-2101

diversity.usc.edu

Information on events, programs and training, the Provost's Diversity and Inclusion Council, Diversity Liaisons for each academic school, chronology, participation, and various resources for students.

USC Emergency - UPC: (213) 740-4321, HSC: (323) 442-1000 – 24/7 on call

dps.usc.edu, emergency.usc.edu

Emergency assistance and avenue to report a crime. Latest updates regarding safety, including ways in which instruction will be continued if an officially declared emergency makes travel to campus infeasible.

USC Department of Public Safety - UPC: (213) 740-6000, HSC: (323) 442-120 – 24/7 on call

dps.usc.edu

Non-emergency assistance or information.