

INFO 523 Syllabus Fall 2025

Course description

INFO 523 Data Mining and Discovery- This course will introduce students to the concepts and techniques of data mining for knowledge discovery. It includes methods developed in the fields of statistics, large-scale data analytics, machine learning, pattern recognition, database technology and artificial intelligence for automatic or semi-automatic analysis of large quantities of data to extract previously unknown patterns. Topics include understanding varieties of data, data preprocessing, classification, association and correlation rule analysis, cluster analysis, outlier detection, and data mining trends and research frontiers. We will use software packages for data mining, explaining the underlying algorithms and their use and limitations. The course include laboratory exercises, with data mining case studies using data from many different resources.

Course Offering

47918-52212 INFO 523 101/201 - Data Mining and Discovery

Instructor Information

John Chen,

Lecturer,

College of Information Science

johnchen@arizona.edu

Office: Harvill 304

Office Hours: Mondays, TBA, Zoom

Prerequisites

Students are assumed to know the basics in computer programming (e.g., variables, arrays, loops, if-then conditions), statistics (e.g., normal distribution, significance tests), and relational database (e.g., ER-diagram, SQL statements). ## Course Format

Asynchronous online lectures: Tuesday, Thursday (Suggested) (At Your Pace), Online, Asynchronous. Set your own schedule to watch the lecture videos, but make sure to turn in your assignments and projects on time.

Course Objective

INFO 523 is a required course in the College of Information Science's M.S. program. As a multidisciplinary field, the course introduce concepts and work from many areas critical to information studies including statistics, machine learning, pattern recognition, database technology, and data visualization.

Learning Outcomes

By the end of this course, students will:

- Understand a large set of concepts of data mining and knowledge discovery.
- Evaluate and use algorithms and software packages to perform data mining analyses.
- Explain and interpret results from data mining analyses.
- (Optional) Exploring the usage of Generative AI in and for data mining.

Textbooks:

- [Data mining conceptual] Jiawei Han, Jian Pei, Hanghang Tong. [Data Mining Concepts and Techniques](#). 4th edition. Morgan Kaufmann, 2023.
- [Python Data Analysis] Wes McKinney. [Python for Data Analysis](#), 3E. O'Reilly, 2023.
- [Prac Stats for DS] Peter Bruce, Andrew Bruce, Peter Gedeck. [Practical Statistics for Data Scientists](#). O'Reilly, 2016.
- [ISL] James Garth, Witten Daniela, Hastie Trevor, Tibshirani Robert. [An Introduction to Statistical Learning](#). Springer, 2021/2023.
- [Pract Time Series] Nielsen Aileen. [Practical Time Series Analysis](#). O'Reilly, 2020.

Course Schedule

An up-to-date schedule, assignments, and due dates can be found on the [course website](#). Dates and time is always in **Mountain Standard Time** (UTC-7).

Course Competencies

The course addresses the MS Competencies: C1 [A, B, C, D], C2, and C3

Course Community

UArizona Community Standard

All students must adhere to the [UArizona Student Rights & Responsibilities](#): The University of Arizona is a community dedicated to scholarship, leadership, and service and to the principles of honesty, fairness, and accountability. Citizens of this community commit to reflect upon these principles in all academic and non-academic endeavors, and to protect and promote a culture of integrity.

Inclusive Community

It is my intent that students from all diverse backgrounds and perspectives be well-served by this course, that students' learning needs be addressed both in and out of class, and that the diversity that the students bring to this class be viewed as a resource, strength, and benefit. It is my intent to present materials and activities that are respectful of diversity and in alignment with [UArizona's Commitment to Diversity and Inclusion](#). Your suggestions are encouraged and appreciated. Please let me know ways to improve the effectiveness of the course for you personally, or for other students or student groups.

Furthermore, I would like to create a learning environment for my students that supports a diversity of thoughts, perspectives and experiences, and honors your identities. To help accomplish this:

- If you have a name that differs from those that appear in your official UArizona records, please let me know! You'll be able to note this in the Getting to know you survey.
- If you feel like your performance in the class is being impacted by your experiences outside of class, please don't hesitate to come and talk with me. If you prefer to speak with someone outside of the course, your academic dean is an excellent resource.
- I (like many people) am still in the process of learning about diverse perspectives and identities. If something was said in class (by anyone) that made you feel uncomfortable, please let me or a member of the teaching team know.

Communication

All lecture notes, assignment instructions, an up-to-date schedule, and other course materials may be found on the course website: <https://info523-fall25-101-201.github.io>.

I will regularly send course announcements via email and Slack, make sure to check one or the other of these regularly. If an announcement is sent Monday through Thursday, I will assume that you have read the announcement by the next day. If an announcement is sent on a Friday or over the weekend, I will assume that you have read it by Monday.

Where to Get Help

- If you have a question during lecture, feel free to ask it! There are likely other students with the same question, so by asking you will create a learning opportunity for everyone.
- The teaching team is here to help you be successful in the course. You are encouraged to attend office hours to ask questions about the course content and assignments. Many questions are most effectively answered as you discuss them with others, so office hours are a valuable resource. Please use them!
- Outside of class and office hours, any general questions about course content or assignments should be posted on the course Slack. There is a chance another student has already asked a similar question, so please check the other posts on Slack before adding a new question. If you know the answer to a question posted on Slack, I encourage you to respond!

Check out the [Support](#) page for more resources.

I want to make sure that you learn everything you were hoping to learn from this class. If this requires flexibility, please don't hesitate to ask.

- You *never* owe me personal information about your health (mental or physical) but you're always welcome to talk to me. If I can't help, I likely know someone who can.
- I want you to learn lots of things from this class, but I primarily want you to stay healthy, balanced, and grounded during this crisis.

Lectures

The goal of the lectures is for them to be as interactive as possible. My role as instructor is to introduce you new tools and techniques, but it is up to you to take them and make use of them. A lot of what you do in this course will involve writing code, and coding is a skill that is best learned by doing. Therefore, as much as possible, you will be working on a variety of tasks and activities throughout each lecture and lab. Attendance will not be taken during

class but you are expected to attend all lecture and lab sessions and meaningfully contribute to in-class exercises and discussion.

You are expected to bring a laptop to each class so that you can take part in the in-class exercises. Please make sure your laptop is fully charged before you come to class as the number of outlets in the classroom will not be sufficient to accommodate everyone. See the [UArizona Libraries loaner technology](#) if you need a loaner laptop.

Assessment

Assessment for the course is comprised of four components: participation, reading quizzes, homework assignments, and projects. All assignments are due by 11:59pm on the indicated day on the [course schedule](#)(