Aimed at extracting useful and interesting patterns and knowledge from large data repositories such as databases and the Web, the field of data mining integrates techniques from databases, statistics, and artificial intelligence. This course will provide a broad overview of the field and focus on a series of advanced topics. The following topics will be covered:

- Knowledge discovery in databases (association rule, clustering, classification);
- Text mining (topic modeling, word embedding, computing journalism);
- Graph mining (PageRank, frequent graph patterns, summarization, linkage prediction);

Students will be required to propose and conduct a course project after the first part (association rule, clustering, classification) is covered. In this project, students will be applying several mining techniques to solve a real-world problem. More details on the project proposal and evaluation standards can be found in the Course Project module. There will be quizzes designed for each topic and weekly assignments.

At the end of the course, students will be able to:

- Understand the process of modeling/transforming a real-world problem to an automatic machine computable knowledge discovery framework;
- Learn the principle of data mining techniques, including association rule mining, clustering, classification;
- Apply clustering/classification techniques to a real-world application, e.g., a movie recommendation system;
- Evaluate the mining outcome using different metrics;
- Apply data mining to text documents, e.g., language model construction, topic modeling, and categorization;
Apply data mining to graph data (e.g., social networks), e.g., computing PageRank and predicting linkage;

PREREQUISITES AND COREQUISITES
The official pre-requisite of the course is COMP 3160. However, as Data Mining is a diverse field, it draws on different aspects of the knowledge in fields such as Databases, Artificial Intelligence, Statistics. The following is a checklist of material that will be used in the course. It is OK if you do not know all of them, but do try to read up on your own.

- Basic computer algorithms (COMP 4030)
- Undergraduate level statistics/probability (ISDS 2710/MATH 4611)
- Database systems (COMP 7115/ISDS 7605)

Programming skill requirement:

- Familiar with any of {C, C++, Java, Python, Matlab}

COURSE TOPICS
Students should finish the first part (association rule mining, clustering, and classification) prior to moving to the text/graph mining part. The first part of this course covers the most fundamental principles for data mining regardless of the underlying data type. While the other two parts can be considered as two important fields of data mining applications in practice.

Textbooks, Supplementary Materials, Hardware and Software Requirements

There is no required textbook for this course. Some recommended but not mandatory extra reading materials are from the following textbook (online digital version available):

- **Data Mining: Concepts and Techniques, 3rd ed.** Jiawei Han, Micheline Kamber, and Jian Pei. Morgan Kaufmann Series in Data Management Systems Morgan Kaufmann Publishers, July 2011. ISBN 978-0123814791


Note that all reading materials will be uploaded to Dropbox for free and easy access.
HARDWARE AND SOFTWARE REQUIREMENTS
The minimum requirements can be found at
https://www.memphis.edu/uofmglobal/services/technology/requirements.php.

Assessment and Grading

ASSESSMENT ASPECTS AND WEIGHTS
As a graduate-level course, the assessment of a student’s performance consists of the following parts:

- Quizzes: 13%
- Weekly assignment: 42%
- Paper review: 10%
- Course Project: (35%)
  - Proposal: 5%
  - Report: 15%
  - Deliverables: 15%

GRADING SCALE
We will calculate final letter grades in two different ways; then each student will receive the higher of the two letter grades. One way is a fixed grading scale, with the following cutoffs:

- A≥90%, A−≥82%, B+≥74%, B≥66%, B−≥58%, C+≥50%, C≥42%

The other way is a curve, with the following percentages of students receiving each grade:

- A: 18%, A−: 18%, B+: 18%, B: 18%, B−: 18%, C+: 5%, C: 5%

Note that extra credits (at most 5%) will be granted for active participation in the course. Any student with truly exceptional performance will be awarded an A+.

Assignments and Participation

ASSIGNMENTS AND PROJECTS
Week 1 Assignment 25 points (3%)
Week 2 Assignment 25 points (3%)

Week 3 Assignment 25 points (3%)

Week 4 Assignment 25 points (3%), Quiz 25 points (4%)

Week 5 Assignment 25 points (3%)

Week 6 Assignment 25 points (3%)

Week 7 Assignment 25 points (3%), Quiz 25 points (3%)

Week 8 Assignment 25 points (3%)

Week 9 Assignment 25 points (3%)

Week 10 Assignment 25 points (3%)

Week 11 Assignment 25 points (3%), Quiz 25 points (3%)

Week 12 Assignment 25 points (3%)

Week 13 Assignment 25 points (3%)

Week 14 Assignment 25 points (3%), Quiz 25 points (3%)

Literature Survey Essay 100 points (10%)

Project Proposal 50 points (5%), Report 100 points (15%), Deliverables 100 points (15%)

Total: 800 Points

CLASS PARTICIPATIONS
Students are expected to communicate with the instructor as a learning resource, students must check the course bulletin board frequently for announcements, and students must actively participate in threaded discussion events.

PUNCTUALITY
All assignments are due before midnight the next Tuesday.

The literature survey essay will be due at 17:00 on Friday of the 8th week.
The course project consists of three parts: a proposal, a report, and deliverables. The proposal will be due at 17:00 on Friday of the 6th week. The deliverables and project report will be due at 17:00 on Friday of the 14th week.

Late submission is allowed but penalties will be applied. Let T be the number of hours after the deadline:

- 5 points deducted if $24 \geq T > 0$;
- 10 points deducted if $48 \geq T > 24$;
- 15 points deducted if $72 \geq T > 48$;
- Submission after 72 hours will not be accepted.

**Course Ground Rules**

- Students are expected to communicate with other students in team projects, learn how to navigate in D2L, and keep abreast of course announcements;
- Students must use the assigned university e-mail address rather than a personal e-mail address;
- Students must observe course netiquette at all times.

**Guidelines for Communication**

**EMAIL**

- Always include a subject line.
- Remember without facial expressions some comments may be taken the wrong way. Be careful in wording your emails. Use of emoticons might be helpful in some cases.
- Use standard fonts.
- Do not send large attachments without permission.
- Special formatting such as centering, audio messages, tables, html, etc. should be avoided unless necessary to complete an assignment or other communication.
- Respect the privacy of other class members.
DISCUSSION GROUPS

- Review the discussion threads thoroughly before entering the discussion. Be a lurker then a discussant.
- Try to maintain threads by using the "Reply" button rather starting a new topic.
- Do not make insulting or inflammatory statements to other members of the discussion group. Be respectful of others' ideas.
- Be patient and read the comments of other group members thoroughly before entering your remarks.
- Be cooperative with group leaders in completing assigned tasks.
- Be positive and constructive in group discussions.
- Respond in a thoughtful and timely manner.

CHAT

- Introduce yourself to the other learners in the chat session.
- Be polite. Choose your words carefully. Do not use derogatory statements.
- Be concise in responding to others in the chat session.
- Be prepared to open the chat session at the scheduled time.
- Be constructive in your comments and suggestion.

WEB RESOURCES

- Columbia Guide to Online Style by Janice R. Walker and Todd Taylor
- Citation Styles Online [http://www.bedfordstmartins.com/online/cite6.html](http://www.bedfordstmartins.com/online/cite6.html)

Plagiarism and Integrity

Plagiarism, cheating, and other forms of academic dishonesty are prohibited. Students guilty of academic misconduct, either directly or indirectly, through participation or assistance, are immediately responsible to the instructor of the class in addition to other possible disciplinary sanctions which may be imposed through the regular institutional disciplinary procedures. Expectations for academic integrity and student conduct are described in detail on the website of
the Office of Student Judicial and Ethical Affairs http://saweb.memphis.edu/judicialaffairs. Please read in particular, the section about "Academic Dishonesty"

**Turnitin.com**

"Your written work will be submitted to Turnitin.com, or a similar electronic detection method, for an evaluation of the originality of your ideas and proper use and attribution of sources. As part of this process, you may be required to submit electronic as well as hard copies of your written work, or be given other instructions to follow. By taking this course, you agree that all assignments may undergo this review process and that the assignment may be included as a source document in Turnitin.com's restricted access database solely for the purpose of detecting plagiarism in such documents. Any assignment not submitted according to the procedures given by the instructor may be penalized or may not be accepted at all."

**Library, Tutoring, and Other Resources**

- The myMemphis Portal system, eCampus Student tab provides access to University library.
- The tutoring link in the course navigation bar provides access to free online tutoring through UpSwing tutoring.
- The Lynda.com link in the course navigation bar provides free access to thousands of video tutorials.
- Other support services are available through the Educational Resources site at: http://www.memphis.edu/students.htm
Students With Disabilities

Qualified students with disabilities will be provided reasonable and necessary academic accommodations if determined eligible by disability services staff at the University of Memphis. Prior to granting disability accommodations in this course, the instructor must receive written verification of a student’s eligibility for specific accommodations from the disability services staff. It is the student’s responsibility to initiate contact with Disability Resources for Students (DRS) and to follow the established procedures for having the accommodation notice sent to the instructor.

Sexual Misconduct and Domestic Violence Policy

This policy specifically addresses sexual misconduct which includes dating violence, domestic violence, sexual assault, and stalking. The policy establishes procedures for responding to Title IX-related allegations of sexual misconduct. Complaints can be reported to the Office for Institutional Equity (OIE). You may contact OIE by phone at 901.678.2713 or by email at oie@memphis.edu. Complaints can be submitted online at File a Complaint. OIE’s office is located at 156 Administration Building.

Non-Discrimination and Anti-Harassment Policy

University policy prohibiting discrimination and harassment based on protected characteristics and classes. Complaints of discrimination and harassment can be reported to the Office for Institutional Equity (OIE). You may contact OIE by phone at 901.678.2713 or by email at oie@memphis.edu. The full text of the policy can be found at GE2030 - Non-Discrimination and Antiharassment.

Technology Requirements

The following is a list of the minimum requirements to use our learning management system. Some courses will have more advanced requirements.

- Access to a reliable, high-speed Internet connection (DSL or Cable).
- Test your device to ensure it is compatible with our LMS (Learning Management System) using the System Check Wizard.
• Open PDF files using the free downloadable PDF software.
• Access Flash-based content with Adobe Flash Player (free).
• Use Microsoft Office for document creation (available for students via umapps.memphis.edu)

Play media content with Real Player (free), QuickTime (free), or Windows Media Player (free).

Syllabus Changes

The instructor reserves the right to make changes as necessary to this syllabus. If changes are necessitated during the term of the course, the instructor will immediately notify students of such changes both by individual email communication and posting both notification and nature of change(s) on the course bulletin board.

Technical Support

Call the Helpdesk: 901-678-8888

Online Helpdesk: https://www.memphis.edu/umtech/service_desk/contact.php