

The University of Texas at El Paso
Department of Computer Science
CS 4364/5364: Information Retrieval

Fall 2023 Syllabus

| | |
|---|--|
| Time and Location: MW 12:00pm-1:20pm CCSB G.0208 | Instructor: Monika Akbar Email: makbar@utep.edu Phone: 915-747-5883 Office: CCSB 3.0422 Office Hours: Tuesday (in-person): 10:30 am to 12:00 pm Wednesday (MS Teams): 4:00 to 5:00 pm Or, by appointment |
|---|--|

Course Objectives: This is an introductory course in information retrieval (IR). This course delves into the theory and practical aspects of information systems, particularly search engines. The curriculum includes topics like text preprocessing, indexing, models and methods for representing text, query processing and optimization, and various services in information systems such as searching, ranking, and recommendation. Through this course, students will build strength in the core areas of information retrieval systems.

Prerequisite: None. Although, it is recommended that the student has taken Data Structures.

Book:

- Introduction to Information Retrieval by Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schütze. 2008, Cambridge University Press, ISBN-10: 0521865719, ISBN-13: 978-0521865715. Available at: <https://nlp.stanford.edu/IR-book/>
- Search Engines: Information Retrieval in Practice. Croft, Metzler and Strohman (Addison-Wesley, 2008). <http://ciir.cs.umass.edu/downloads/SEIRiP.pdf>
- Reference books:
 - Web information retrieval by Stefano Ceri, Alessandro Bozzon, Marco Brambilla, Emanuele Della Valle, Piero Fraternali, Silvia Quarteroni. Heidelberg. Springer, 2013. (E-book free through the UTEP library.)
- Additional resources will be available in Blackboard.

Course Website: Blackboard.

GRADING

Grades are communicated to students in a timely manner. It is the students' responsibility to keep track of their grades. The semester grade will be based on a combination of homework assignments, quizzes, class participation, two midterm exams, and a project. Graduate students will have additional sections in some assignments. Additionally, graduate students will also have presentations. The approximate percentages of these activities are as follows:

- Quiz and in-class participation: 20%
- Homework/assignments: 20%
- Midterm 1: 20%
- Midterm 2: 20%
- Project: 20%

The nominal percentage-score-to-letter-grade conversion is as follows:

- 90% or higher is an A
- 80-89% is a B
- 70-79% is a C
- 60-69% is a D
- Below 60% is an F

Graduate students: Graduate students will have additional sections in assignments. Additionally, graduate students will also have presentations for the project.

Grade appeals: Any grades must be appealed within seven (7) days of the grade being posted.

COURSE OUTLINE

Class Description: This class will include lectures, in-class activities, state-of-the-art literature surveys, and a semester-long group project. The class will include presentations by the instructor and students, discussions, and demos of various concepts. Students will learn to think critically, communicate effectively, and collaborate productively in a group.

Tentative Weekly Schedule:

| Week | Topic |
|------|--|
| 1 | Introduction to information retrieval: history, key concepts, and terminology. |
| 2 | Transformation: Text preprocessing and indexing |
| 3 | Retrieval models – Boolean, vector space, weighted vector space |
| 4 | Probabilistic model and language models. |
| 5 | Evaluation of IR tasks |
| 6 | Midterm 1 |
| 7 | Retrieval: searching, scoring, and ranking of search results. Topic analysis. Query processing and optimization. |
| 8 | Web search engine |
| 9 | Graph-based information retrieval |
| 10 | Distributed word representations |
| 11 | Large language models: BERT and GPT-based models for IR tasks |
| 12 | Midterm 2 |
| 13 | Project presentation |
| 14 | Project presentation |
| 15 | Review and applications of information retrieval: recommendation, filtering |

COURSE DETAILS

Exams: There will be **two midterm exams**. There will be no final exam. The two midterm scores will weigh 40% of your overall class grade. Make-up exams will not be permitted except under unusual circumstances with satisfactory written justification. Any student who misses an exam due to an unexcused absence will receive a grade of zero for that exam with no opportunity for make-up or substitution. University excused absences will be excused; the exam-related arrangements should be made in advance in those cases.

Homework: Regular homework/assignments will be assigned, which will require significant effort outside of class. The assignments are designed to challenge you by requiring that you apply learned concepts to new situations. You should start working on your assignment immediately after receiving it.

Quizzes and Exercises: There will be regular quizzes and exercises in the class. The purpose of each quiz and exercise is to ensure that you are staying current with the class content and to verify that you have acquired the skills introduced in the class. The quizzes are not pre-scheduled. There will be **no make-up** on missed quizzes. Please stay up-to-date with class content and activities to perform well in the quiz.

Project: There will be an individual project for the graduate and undergraduate students. The instructor will provide a list of potential projects that students can choose from. A student can create his/her own project, too, with approval from the instructor.

The project must have information retrieval components. The target of the project is to deliver a publishable report and presentation (only for graduate students).

Class Participation: Students should be **on time** for all scheduled sessions and **attend the entire session**. Attendance at and participation in all lecture sessions are essential for your success in this course. Although attendance has a specific weight, the instructor reserves the right to penalize the final grade for low attendance because active in-class participation is critical in this class. Taking notes on paper is recommended than using digital devices.

Late Policy: All assignments are due before the start of class on the day it is due, unless otherwise stated. There will be a 25% penalty for each day after the deadline. No submission will be accepted after the 4th day.

RESOURCES

Accommodations Policy: UTEP is committed to providing reasonable accommodations and auxiliary services to students, staff, faculty, job applicants, applicants for admissions, and other beneficiaries of University programs, services and activities with documented disabilities in order to provide them with equal opportunities to participate in programs, services, and activities in compliance with sections 503 and 504 of the Rehabilitation Act of 1973, as amended, and the Americans with Disabilities Act (ADA) of 1990 and the Americans with Disabilities Act Amendments Act (ADAAA) of 2008. Reasonable accommodations will be made unless it is determined that doing so would cause undue hardship on the University. Students requesting an accommodation based on a disability must register with the UTEP [Center for Accommodations and Support Services](#) (CASS); please contact the office at (915) 747-5148, or by email to cass@utep.edu. Students are required to discuss their accommodations with the instructor for a proper plan to be made.

Help Desk: Please contact UTEP [Help Desk](#) if you are experiencing technological challenges (email, Blackboard, software, etc.), and submit a digital ticket for assistance. Contact the Helpdesk via phone, email, chat, website, or in-person if on campus.

Individual Resources

- [Military Student Success Center](#): Assists personnel in any branch of service.
- [Counseling and Psychological Services](#): Provides a variety of counseling services, including individual, couples, and group sessions as well as career and disability assessments.

Standards of Conduct

You are expected to conduct yourself in a professional and courteous manner, as prescribed by the UTEP [Standards of Conduct](#).

A fundamental principle for any educational institution, academic integrity is highly valued and seriously regarded at The University of Texas at El Paso. More specifically, students are expected to maintain absolute integrity and a high standard of individual honor in scholastic work undertaken at the University. At a minimum, you should complete any assignments, exams, and other scholastic endeavors with the utmost honesty, which requires you to:

- Acknowledge the contributions of other sources to your scholastic efforts.
- Complete your assignments independently unless expressly authorized to seek or obtain assistance in preparing them.
- Follow instructions for assignments and exams, and observe the standards of your academic discipline; and
- Avoid engaging in any form of academic dishonesty on behalf of yourself or another student.

Graded work, e.g., homework and tests, is to be completed independently and should be unmistakably your own work (or, in the case of group work, your team's work), although you may discuss your project with other students in a general way. You may not represent as your own work material that is transcribed or copied from another person, book, or any other source, e.g., a web page.

Any student who commits an act of scholastic dishonesty is subject to discipline. Scholastic dishonesty includes, but not limited to cheating, plagiarism, collusion, the submission for credit of any work or materials that are attributable to another person.

Scholastic Dishonesty: Any student who commits an act of scholastic dishonesty is subject to discipline. Scholastic dishonesty includes, but not limited to cheating, plagiarism, collusion, submission for credit of any work or materials that are attributable to another person.

Cheating is copying from the test paper of another student. Communicating with another student during a test to be taken individually. Giving or seeking aid from another student during a test to be taken individually. Possession and/or use of unauthorized materials during tests (i.e., crib notes, class notes, books, etc.). Substituting for another person to take a test. Falsifying research data, reports, academic work offered for credit.

Plagiarism is using someone's work in your assignments without the proper citations. Submitting the same paper or assignment from a different course, without direct permission of instructors.

To avoid plagiarism, see: <https://www.utep.edu/student-affairs/osccr/Files/docs/Avoiding-Plagiarism.pdf>

Plagiarism Detection: All coursework and assignments are subject to be submitted to plagiarism detection software including, but not limited to SafeAssign.

Collusion is unauthorized collaboration with another person in preparing academic assignments.

Using generative AI: Exercise caution when utilizing AI tools like ChatGPT. It is acceptable to use these tools to complement learning, but their use is restricted during exams and quizzes. Students should work on developing their proficiency in applying introduced concepts. Proper citation of sources is required to adhere to citation guidelines and copyright regulations, and to prevent plagiarism when utilizing AI tools. Violations of these guidelines will lead to referrals to the Office of Student Conduct and Conflict Resolution.

NOTE: When in doubt about any of the above, please contact your instructor to check if you are following the authorized procedure.

COVID-19 PRECAUTIONS

Please stay home if you have been diagnosed with COVID-19 or are experiencing COVID-19 symptoms. If you are feeling unwell, please let me know as soon as possible, so that we can work on appropriate accommodations. If you have tested positive for COVID-19, you are encouraged to report your results to covidaction@utep.edu, so that the Dean of Students Office can provide you with support and help with communication with your professors. The Student Health Center is equipped to provide COVID 19 testing.

The Center for Disease Control and Prevention recommends that people in areas of substantial or high COVID-19 transmission wear face masks when indoors in groups of people. The best way that Miners can take care of Miners is to get the vaccine. If you still need the vaccine, it is widely available in the El Paso area, and will be available at no charge on campus during the first week of classes. For more information about the current rates, testing, and vaccinations, please visit epstrong.org