The University of Texas at El Paso
Department of Computer Science
CS 4390/5390: Information Retrieval and Information Visualization

Spring 2020 Syllabus

Time and Location: MW 12:00 PM – 1:20 PM CCSB 1.0410	**Instructor:** Monika Akbar **Email:** makbar@utep.edu **Phone:** 915-747-5883 **Office:** CCSB 3.0422 **Office Hours:** TBD

Course Objectives: This is an introductory course in information retrieval (IR) and information visualization. IR covers theory and practice of information systems such as search engines. Topics in IR include models and methods for representing text, indexing, browsing, and various services in information systems such as searching and recommendation. Through this course, students will learn the fundamentals of information retrieval systems.

Students will also gain knowledge on information visualization principles. Students will be introduced to various information visualization techniques enabling them to choose appropriate visualizations for different types data and for different tasks.

Prerequisite: None. Although, it is recommended that 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
- Interactive Data Visualization for the Web by Scott Murray. O'Reilly Media, Inc. Print ISBN-13: 978-1-4919-2128-9. (E-book free through the UTEP library.)
- 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.)
    - The Visual Display of Quantitative Information by Edward R Tufte.
- 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. Your semester grade will be based on a combination of homework assignments, quizzes, class participation, 2 midterm exams, a final exam, and a project. The approximate percentages are as follows:

- Quiz and in-class participation: 20%
- Homework/assignments: 20%

- Exams: 30%
- Project: 30%

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

Grade appeals: Any grades, except the final exam, must be appealed within 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 approaches introduced throughout the semester. Students will learn to think critically, communicate effectively, and collaborate productively in a group.

Tentative Weekly Schedule:

Week	Topic
1	Introduction to Information Retrieval (IR). Introduction to data properties: data architecture, data types and data formats, data modeling and design, linked open data.
2	Introduction to IR tasks: Acquisition (crawling), text pre-processing., indexing (construction and compression), retrieval, query processing.
3	Transformation: Information representation - local and distributed representations, weighted vector-space model.
4	Retrieval models: Probabilistic graphical model, Latent semantic indexing. Retrieval evaluation: Scoring and ranking of search results.
5	Midterm 1
6	IR Applications: Filtering, recommendation, personalization. Introduction to Information Visualization.
7	Design Principles: HCI metrics, scale, tasks, data mapping, overviewing, data density, graphic properties, presentation methods.
8	Visual Encoding: Graphical methods for data presentation, visual perception, graphical perception, layering and separation.
9	Interaction Strategies: Overviews, multiple views, focus + context, filtering and querying, lens, data flow.
10	Midterm 2
11	Introduction to D3 visualization.
12	Applications of visualizations: Explore, transform, visualize large complex datasets including scientific data, tabular data, text data, graphs, and documents.
13	Project presentation
14	Review
15	Final exam

COURSE DETAILS

Project: There will be a semester long group project for the graduate and undergraduate students. Students are encouraged to form groups of three to four members --- including both graduate and undergraduate students. The instructor will provide a list of potential projects that teams can choose from. A team can create their own project too, with approval from the instructor.

The project must have information retrieval and visualization components. The target of the project is to deliver a publishable report along with the description of the methodology and full experimental results.

Exams: There will be **2 midterm exams** and **one final exam**. All three exams together will weigh 30% of your overall final 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/assignment 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.

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 4^{th} day.

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.

Use of Electric Devices in Class: You are requested to appropriately use any digital device in class. Cell phones should be turned off or put in 'silent' mode in the class. Please do not read text messages or send text messages during the class.

It is recommended that you bring a laptop in the classroom, however, please do not chat, check email, browse social media, surf the web, etc. unless the instructor advises you to do so. Such activities distract you from learning as well as affect those sitting next to you.

Taking notes on paper is recommended than using digital devices.

Wearing headphones of any type are strictly prohibited in the classroom. Headphones can be permitted only if the student has appropriate documentation approved by the Center for Accommodations and Support Services (CASS).

RESOURCES

Special Accommodations: If you have a disability and need classroom accommodations, please contact the Center for Accommodations and Support Services (CASS) at 747-5148 or by email to cass@utep.edu, or visit their office located in the UTEP Union East, Room 106. For additional information, please visit the CASS website at www.sa.utep.edu/cass. CASS' staff are the only individuals who can validate and if need be, authorize accommodations for students with disabilities.

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

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

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