

Course Information

Course Number: CSCE 670
Course Title: Information Storage and Retrieval
Section: 600
Time: Tuesdays and Thursdays 5:30pm – 6:45pm
Location: HRBB 113
Credit Hours: 3

Instructor Details

Instructor: Yu Zhang
Office: PETR 222
E-Mail: yuzhang@tamu.edu
Office Hours: Thursdays 4pm – 5pm

Course Description

In this course, we will study the theory, design, and implementation of foundational information retrieval (IR) systems, but also examine closely modern web search and recommender systems, including algorithms and techniques at the core of how people connect to information. Broadly, what are the principal ideas, algorithms, and systems for organizing information?

Course Prerequisites

CSCE 310 or CSCE 603 or approval of instructor; graduate classification.

Course Learning Outcomes

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

- Define and explain the key concepts and models relevant to web search, including topics like text indexing, retrieval models, evaluation, Web crawling, link-based algorithms like PageRank, and learning to rank.
- Define and explain the key concepts and models relevant to recommender systems, including topics like collaborative filtering, matrix factorization, recommender system evaluation, and implicit recommendation.
- Design, implement, and evaluate the core algorithms underlying a fully functional web search system and recommendation system.
- Identify the salient features and apply recent research results in web search and recommender systems, including topics such as adversarial information retrieval and neural models of retrieval and recommendation.

Textbook and/or Resource Materials

Required Textbooks

This course does not mandate any textbook. The lecture slides/videos and other materials provided by the instructor will be sufficient, serving as the primary reference. In addition, the students will read some selections from:

- **(MRS) *Introduction to Information Retrieval***, Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schutze, 2009.
 - Available online: <https://nlp.stanford.edu/IR-book/>
- **(LRU) *Mining of Massive Datasets***, Jure Leskovec, Anand Rajaraman, and Jeffrey D. Ullman, 2014.
 - Available online: <http://www.mmds.org/>

Grading Policy

The course grading scale is:

A : $\geq 90\%$

B : $\geq 80\%$ and $< 90\%$

C : $\geq 70\%$ and $< 80\%$

D : $\geq 60\%$ and $< 70\%$

F : $< 60\%$

Homework 30%: We will have five programming assignments (2%+7%+7%+7%+7%). All of them will be in Python. The first assignment is an introductory one designed to help you get familiar with Jupyter Notebook and the Python operations commonly used in this course.

Homework collaboration policy: Homeworks are individual assignments. That is, you should write your own code, by yourself. However, we do encourage you to talk amongst yourselves about approaches and methods. As an example, it is appropriate to sit in a group with no laptops to talk through the problems. It is inappropriate, however, to pair program, share code, or work through specific problems in code together. You must acknowledge all help in your homework submission.

Homework late day policy: All homeworks are due by 11:59pm on the due date. Each student gets five late days total. A late day is an indivisible 24-hour unit. For example, if the due date is 11:59pm on Monday, and you submit at 12:01am Tuesday, that is 1 late day. Once you use up the five late days, a late assignment will receive a grade of 0.

Computer Code. In principle you may submit AI-generated code, or code that is based on or derived from AI-generated code, as long as this use is properly documented in the comments: you need to include the prompt and the significant parts of the response. AI tools may help you avoid syntax errors, but there is no guarantee that the generated code is correct. It is your responsibility to identify errors in program logic through comprehensive, documented testing. Moreover, generated code, even if syntactically correct, may have significant scope for improvement, in particular regarding separation of concerns and avoiding repetitions. The submission itself must meet our standards of attribution and validation.

Quizzes 20%: We will have four brief in-class quizzes.

Final 25%: We will hold a comprehensive final exam from 3:30pm – 5:30pm on Dec 16 in HRBB 113.

Group Project 20%: Students will work in teams of 3 or 4 to engage in a mini-research project.

Late Work Policy

With exception of excused absences allowed by student rule 7, the only late work policy is for homeworks as discussed above.

Work submitted by a student as makeup work for an excused absence is not considered late work and is exempted from the late work policy (Student Rule 7 (<https://student-rules.tamu.edu/rule07/>)).

Regrade Policy: Once you receive your graded assignment (e.g., a homework assignment), you have seven days to request a regrade in writing. After seven days, there will be no regrades. You must give the instructor a written explanation of your regrade request. We reserve the right to regrade the entire assignment.

Course Schedule

(this is a rough schedule for the course; subject to change)

Week	Topics
1	Course Overview; Boolean Retrieval
2	Vector Space Model; TF-IDF; BM25
3	Link Analysis
4	Evaluation of IR Models
5	Learning to Rank
6	Recommender Systems (Collaborative Filtering & Content-Based Approach)
7	Recommender Systems (Matrix Factorization & Bayesian Personalized Ranking)
8	Word Embedding; Neural Ranking
9	Contextualized Language Models; BERT-Based Ranking
10	Neural Collaborative Filtering; Sequential Recommendation
11	Large Language Models Basics
12	Large Language Models for Ranking
13	Large Language Models for Recommendation
14	Information Retrieval for Science and Scientific Research
15	Final Project Presentations

Important Dates (Dates are specific to Fall 2025)

- Homeworks:
 - Homework 0: due Sep 6
 - Homework 1: due Sep 20
 - Homework 2: due Oct 11
 - Homework 3: due Nov 1
 - Homework 4: due Nov 22
-

- Quizzes:
 - Quiz 1: in the Sep 18 class
 - Quiz 2: in the Oct 7 class
 - Quiz 3: in the Oct 30 class
 - Quiz 4: in the Nov 20 class
- Group Project:
 - Proposal: due Oct 18
 - Presentation: in the Dec 2 and Dec 4 classes
 - Report: due Dec 9
- Final: 3:30pm – 5:30pm, Dec 16

University Policies

Attendance Policy

The university views class attendance and participation as an individual student responsibility. Students are expected to attend class and to complete all assignments.

Please refer to [Student Rule 7](#) in its entirety for information about excused absences, including definitions, and related documentation and timelines.

Makeup Work Policy

Students will be excused from attending class on the day of a graded activity or when attendance contributes to a student's grade, for the reasons stated in Student Rule 7, or other reason deemed appropriate by the instructor.

Please refer to [Student Rule 7](#) in its entirety for information about makeup work, including definitions, and related documentation and timelines.

Absences related to Title IX of the Education Amendments of 1972 may necessitate a period of more than 30 days for make-up work, and the timeframe for make-up work should be agreed upon by the student and instructor" ([Student Rule 7, Section 7.4.1](#)).

"The instructor is under no obligation to provide an opportunity for the student to make up work missed because of an unexcused absence" ([Student Rule 7, Section 7.4.2](#)).

Students who request an excused absence are expected to uphold the Aggie Honor Code and Student Conduct Code. (See [Student Rule 24](#).)

Academic Integrity Statement and Policy

"An Aggie does not lie, cheat or steal, or tolerate those who do."

“Texas A&M University students are responsible for authenticating all work submitted to an instructor. If asked, students must be able to produce proof that the item submitted is indeed the work of that student. Students must keep appropriate records at all times. The inability to authenticate one’s work, should the instructor request it, may be sufficient grounds to initiate an academic misconduct case” ([Section 20.1.2.3, Student Rule 20](#)).

You can learn more about the Aggie Honor System Office Rules and Procedures, academic integrity, and your rights and responsibilities at aggiehonor.tamu.edu.

Americans with Disabilities Act (ADA) Policy

Texas A&M University is committed to providing equitable access to learning opportunities for all students. If you experience barriers to your education due to a disability or think you may have a disability, please contact the Disability Resources office on your campus (resources listed below). Disabilities may include, but are not limited to attentional, learning, mental health, sensory, physical, or chronic health conditions. All students are encouraged to discuss their disability related needs with Disability Resources and their instructors as soon as possible.

Disability Resources is located in the Student Services Building or at (979) 845-1637 or visit disability.tamu.edu.

Title IX and Statement on Limits to Confidentiality

Texas A&M University is committed to fostering a learning environment that is safe and productive for all. University policies and federal and state laws prohibit gender-based discrimination and sexual harassment, including sexual assault, sexual exploitation, domestic violence, dating violence, and stalking.

With the exception of some medical and mental health providers, all university employees (including full and part-time faculty, staff, paid graduate assistants, student workers, etc.) are Mandatory Reporters and must report to the Title IX Office if the employee experiences, observes, or becomes aware of an incident that meets the following conditions (see [University Rule 08.01.01.M1](#)):

- The incident is reasonably believed to be discrimination or harassment.
- The incident is alleged to have been committed by or against a person who, at the time of the incident, was (1) a student enrolled at the University or (2) an employee of the University.

Mandatory Reporters must file a report regardless of how the information comes to their attention – including but not limited to face-to-face conversations, a written class assignment or paper, class discussion, email, text, or social media post. Although Mandatory Reporters must file a report, in most instances, a person who is subjected to the alleged conduct will be able to control how the report is handled, including whether or not to pursue a formal investigation. The University’s goal is to make sure you are aware of the range of options available to you and to ensure access to the resources you need.

Students wishing to discuss concerns in a confidential setting are encouraged to make an appointment with [Counseling and Psychological Services](#) (CAPS).

Students can learn more about filing a report, accessing supportive resources, and navigating the Title IX investigation and resolution process on the University’s [Title IX webpage](#).

Statement on Mental Health and Wellness

Texas A&M University recognizes that mental health and wellness are critical factors that influence a student's academic success and overall wellbeing. Students are encouraged to engage in healthy self-care by utilizing available resources and services on your campus.

Students who need someone to talk to can contact Counseling & Psychological Services (CAPS) or call the TAMU Helpline (979-845-2700) from 4:00 p.m. to 8:00 a.m. weekdays and 24 hours on weekends. 24-hour emergency help is also available through the 988 Suicide & Crisis Lifeline (988) or at [988lifeline.org](https://www.988lifeline.org)