CS734/834: Introduction to Information Retrieval (Fall 2025)

Instructor

Jian Wu

Email

jwu@cs.odu.edu

Classroom

DRGS 2106

Virtual Office Hours

4 – 5 pm Tuesday or by appointment

Class Time

6 pm. - 8:40 pm. Tuesday

Important Dates

Tuesday, 8/26/2025: first class

Tuesday 12/9/2025: last class

Course Overview

Information retrieval (IR) is the process for a computer system to respond to a user's query based on a collection of information. The IR theory laid the foundation of online search engines. IR was one of the first and remains one of the important problems in computer science. This class will explore the theory and practice of information retrieval in the context of developing web-based search applications and exploring solutions to challenging IR problems. The course will explore issues related to web crawling, ranking, query processing, retrieval models, evaluation, clustering, text classification, and re-ranking. The course will also cover recently established algorithms on query understanding, learning to rank, and neural ranking models. The class will feature hands-on development and coding ElasticSearch, as well as applications of machine learning to IR problems.

Course Delivery Method

This course will be held live (face to face) for the whole semester, unless otherwise be changed per university policies. **In-person attendance is expected during the class time.**

Required Text

There is no required textbook. The recommended textbooks are

Introduction to Information Retrieval by Christopher D. Manning, Prabhakar Raghavan and Hinrich Schutze, Cambridge University Press. 2008, ISBN-13: 978-0521865715, ISBN-10: 0521865719 (hardcover). The online version is here. We will call this book *IIR*.

Search Engines - Information Retrieval in Practice, W. B. Croft, D. Metzler, and T. Strohman. Cambridge University Press. 2009. A free PDF version is here.

Hardware and Software Requirements

Students will need frequent access to a PC (with Windows 10+) or a Mac (with MacOS 10.14+) capable of hosting application development activities or of connecting to remote servers.

The students have options to work on their projects on a virtual environment hosted by the ODU Computer Science. The course will introduce students to a wide variety

of open-source software packages. Students will need to install some of these on their assigned virtual machine.

Alternatively, students can use their own computers or ODU's HPC cluster (Wahab or Turing) for course projects. It should be noted that Wahab and Turing provides GPU access, but they do not provide remote desktop services.

Course Materials

- Course materials and other resources including slides and assignments will be distributed as the course proceeds in the semester.

Grading Policy

Students are graded based on the following aspects.

- Attendance: 10%
- Discussion using Packback: 10%
- Homework: 30%
- Project: 50%

Grading Chart

A	**A-**	**B+**	**B**	**B-**	**C+**	**C***
94-100	90-93.99	87-89.99	84-86.99	80-83.99	77-79.99	74-76.99

* A provisional graduate student who receives a cumulative GPA less than 3.0 will be subject to removal from the graduate program.

Attendance Policy

In-person attendance is required. One absence causes a deduction of 1% on attendance until all points are deducted in this aspect. If more than 10 absences are observed, the student automatically gets an F for this course. In case of absence due to legitimate reasons, including but not limited to sickness, university-approved curricular and extracurricular activities (such as athletic contests), career interviews, the death of family members, students should be prepared to provide documentation at least one day **before** classes. Makeup classes are not available, but students can meet with the instructor in office hours.

Academic Integrity

Individual assignments must be completed **independently**. Students are encouraged to form study groups or to learn from their peers. However, discussion should be limited to **general** approaches. **Specific answers should never be discussed**. ODU's policy regarding Academic Integrity must be followed.

- **Cheating**: Using unauthorized assistance, materials, study aids, or other information in any academic exercise (Examples of cheating include, but are not limited to, the following: using unapproved resources or assistance to complete an assignment, paper, project, quiz or exam; collaborating in violation of a faculty member's instructions; and submitting the same, or substantially the same, paper to more than one course for academic credit without first obtaining the approval of faculty).
- **Plagiarism**: Using someone else's language, ideas, or other original material without acknowledging its source in any academic exercise. 4 Examples of plagiarism include, but are not limited to submitting a research paper obtained from a commercial research service, the Internet, or from another student as if it were original work; or making simple changes to borrowed materials while leaving the organization, content, or phraseology intact. Plagiarism also occurs in a group project if one or more of the members of the group does none of the group's work and participates in none of the group's activities but attempts to take credit for the work of the group.
- **Fabrication**: Inventing, altering or falsifying any data, citation or information in any academic exercise. Examples of fabrication include, but are not limited to, the following: citation of a primary source which the student actually obtained from a secondary source; or invention or alteration of experimental data without appropriate documentation (such as statistical outliers).
- **Facilitation**: Helping another student commit, or attempt to commit, any Academic Integrity violation, or failure to report suspected Academic Integrity violations to a faculty member. An example of facilitation may include circulating course materials when the faculty member has not explicitly authorized their use.

Copyright

- All course materials students receive or to which students have online access are protected by copyright. Students may use course materials and make copies for their own use as needed, but **unauthorized distribution and/or uploading of materials without the instructor’s express permission is strictly prohibited**.

Disability Accommodations

- In order to receive consideration for reasonable accommodations, you must contact the appropriate services office will provide you with an accommodation letter. Please share this letter with your instructors and discuss the accommodations with them as early in your courses as possible. The detail of disability accommodations is documented in ODU policy #4500.

Discrimination and Harassment

- The university is committed to equal access to programs, facilities, admission and employment for all persons. It is the policy of the university to maintain an environment free of harassment and free of discrimination against any person because of age, race, color, ancestry, national origin, religion, creed, service in the uniformed services (as defined in state and federal law), veteran status, sex, sexual orientation, marital or family status, pregnancy, pregnancy-related conditions, physical or mental disability, gender, perceived gender, gender identity, genetic information or political ideas. Discriminatory conduct and harassment, as well as sexual misconduct and relationship violence, violates the dignity of individuals, impedes the realization of the university's educational mission, and will not be tolerated.
- Gender-based sexual harassment, including sexual violence, are forms of gender discrimination in that they deny or limit an individual's ability to participate in or benefit from University programs or activities. These policies shall not be construed to restrict academic freedom at the university, nor shall they be construed to restrict constitutionally protected expression. The policy is coded in University Policy #1005.

Course Schedule*

Week	Dates	Subject	Reading and Homework
1	Tuesday, 8/26/2025	Course Introduction, Projects	
		Internet and Search Engines	IIR Chapters 1, 8.
2	Tuesday, 9/2/2025	Ad-hoc Retrieval and Scalability	IIR Chapter 8
		Ad-hoc Retrieval Evaluations	Homework Set 1
3	Tuesday, 9/9/2025	Inverted index	IIR Chapters 4, 6
		Project proposal presentation	
4	Tuesday, 9/16/2025	Querying processing	Homework Set 2
		Text processing	IIR Chapters 3,5 Homework Set 1 due

Week	Dates	Subject	Reading and Homework
5	Tuesday, 9/23/2025	Traveling, No Class	
6	Tuesday, 9/30/2025	Classification Models	IIR Chapters 13, 14, and 15
		Text Classification	IIR Chapters 13, 14, and 15
7	Tuesday, 10/7/2025	Text representation and Vector Space Model	IIR Chapter 6
		Probability model: BM25, BM25F	IIR Chapter 11
8	Tuesday, 10/14/2025	Fall Holiday, No Class	
9	Tuesday, 10/21/2025	Index Compression	IIR Chapter 5 Homework Set 2 due, Homework Set 3
		Spell Correction	IIR Chapter 3
10	Tuesday, 10/28/2025	Mid-term project presentation	
11	Tuesday, 11/4/2025	Election Day, No Class	
12	Tuesday, 11/11/2025	Web crawling	IIR Chapters 6, 7, 19 & 20
		Link Analysis	IIR Chapter 21 Homework Set 3 due
13	Tuesday, 11/18/2025	Distributed word representations	Mitra (2018); Mikolov et al. (2013); Devlin et al. (2019)
		Vector search	
14	Tuesday, 11/25/2025	Learn to Rank	IIR Sections 6.1 & 15.4 Liu (2014)
		IR and LLM	
15	Tuesday, 12/2/2025	Invited Speaker (TBD)	
		Project time	
16	Tuesday	Project presentation	
	12/9/2025	Project presentation	Final report due by midnight

* Course schedules are subject to change depending on availability of speakers and the instructor.

Exam Schedule

No final exams.