

# CS824 Information Retrieval

## Syllabus (Spring 2021)

### Course information

- **Instructor:** Dr. Mengjun Hu
- **Email:** Course Email at UR Courses (preferred); Mengjun.Hu@uregina.ca
- **Lecture hours and locations:**

Tuesdays, Thursdays 9:00-10:15pm via Zoom meetings (May 3, 2021 - June 16, 2021)

The Zoom link is posted to UR Courses. When joining the Zoom meeting, you should use your real name, including both first and last names.

- **Office hours:** By appointment only.

### Lecture delivery

The lecture videos will be pre-recorded and uploaded to UR Courses at least one day before the scheduled lecture time. The lecture time will be used as Q&A sessions, held via Zoom meetings. Questions about the content presented in the corresponding videos will be given preference. Once such questions are done, you may ask other questions regarding this course. The Q&A sessions will not be recorded due to privacy issues. Attending Q&A sessions is not mandatory, but you are supposed to be available during the scheduled lecture time. Attendance will not be checked for any lecture, but certain participation during Q&A sessions may be considered as evidence when judging for instructor discretion (see more details in Grading and evaluations). Remember this is a compressed semester. For each week, you are expected to spend twice of the time you would usually spend on a course in a normal semester. Watch lecture videos and learn timely. Otherwise, you will easily and quickly fall behind with the lectures. Do your assignments gradually as you learn. Do not leave them at the last minute.

### Topics

Copied from CS website: “Content analysis; types of storage and retrieval systems; retrieval models; information theory; multimedia retrieval; hypertext; information network and inference. Prior to registering for this course, students should have a background in algorithms and data structures, and database and information retrieval comparable to the senior undergraduate level.”

A list of tentative topics includes:

- Basic concepts of information retrieval systems
- Indexing theory
- Boolean retrieval model
- Vector space model
- Probability distribution model
- Probabilistic models
- System evaluation
- Relevance feedback

## Grading and evaluations

- Assignments<sup>1</sup> (5) 30%
- Midterms<sup>2</sup> (2) 30%
- Final exam<sup>3</sup> 40%
- Instructor discretion<sup>4</sup> ±5%

<sup>1</sup> All assignments should be submitted through UR Courses. The accepted file format is pdf. If you use any handwriting app on tablets, you should make sure the generated pdf file could be properly opened by Adobe Reader. Otherwise, you are responsible for any part truncated and any mark lost due to it. No late submission is accepted. (Late for a couple of minutes is fine.) There is no penalty policy for late submissions which are simply not accepted. Add proper references to avoid academic plagiarism whenever appropriate, especially for open-style questions. Use your own words to express your own ideas. Copying from anywhere, either exact content or with simple literal rephrasing, is serious academic misconduct. Any academic misconduct will be reported, and you will get a preliminary penalty of 0 mark while waiting for formal investigation results.

<sup>2</sup> The two midterms will be open book and held through UR Courses. The time and navigation will be restricted. Please be aware that consulting resources are permitted for open-book exams, but not copying. Copying from anywhere, either exact content or with simple literal rephrasing, is serious academic misconduct in any form of exams. You are not permitted to post any part of the exams to anywhere. Such postings are also considered as academic misconduct and will also be reported. You must complete the exams by yourself. Use your own words to express your own ideas. Any academic misconduct will be reported, and you will get a preliminary penalty of 0 mark while waiting for formal investigation results. Makeups for midterms will NOT be offered for any personal reason. Makeups will only be considered in situations that affect most or all students (such as failure of UR Courses quiz functionalities). The exam links will be posted to UR Courses in advance, with details in the descriptions. The midterms are scheduled as follows:

- Midterm 1: May 18, 2021, Sections 1-3.
- Midterm 2: June 8, 2021, Sections 4-6.

The coverage will also be marked on UR Courses. The instructor be available through Zoom during exam time.

<sup>3</sup> The final examination will be held through UR Courses on June 21, 2021 starting from 9:00am (as scheduled). The final examination may be shortened which means it may end before 5:00pm. The final examination will be **closed book** and will be proctored by ProctorTrack. You must make sure your devices meet the requirements for ProctorTrack (Further information can be found here: <https://www.uregina.ca/remote-learning/index.html#proctoring>). Any academic misconduct will be reported and you will get NR mark as your final grade for the course while waiting for final investigation results, which may result in month-long restriction on your registration.

<sup>4</sup> Instructor discretion will only be used to adjust final grades so that they truthfully (as much as possible) reflect how much the students have learned, based on what the students **have shown to the instructor**. Thus, certain participation during lecture time (e.g., asking questions related to lecture content) may be considered when judging for instructor discretion. This doesn't mean that participation will necessarily lead to instructor discretion awarded. The instructor discretion will not be used for simply making a student pass the course.

## Textbook

C.J. van Rijsbergen, Information Retrieval

[http://openlib.org/home/krichel/courses/lis618/readings/rijsbergen79\\_infor\\_retriev.pdf](http://openlib.org/home/krichel/courses/lis618/readings/rijsbergen79_infor_retriev.pdf)

<http://www.dcs.gla.ac.uk/Keith/Preface.html>

A list of additional papers will be posted on UR Courses.

## Additional reference (optional)

R. Baeza-Yates and B. Ribeiro-Neto, Modern Information Retrieval, Addison Wesley, New York, 1999

## Further information

- Your UofR Webmail account is the only official email account for communications regarding this course. You should always use your UofR Webmail account to send emails to the instructor at Mengjun.Hu@uregina.ca, or use Course Email on UR Courses. Otherwise, it will be your responsibility if the emails are not successfully delivered to the instructor's inbox (e.g. identified as junk mails). No excuse will be accepted. You should follow additional instructions if there is any for specific course works such as exams.
- No slide will be used. You should take your own notes when watching lecture videos.
- Foreseen situations (e.g. sickness) that may affect midterms and final should be reported to the instructor before the midterms and final. Unforeseen situations (e.g. loss of power or internet on the days of midterms or final) should be reported to the instructor within the same day of the midterms or final (cutoff time at 23:59pm of the day). Otherwise, you will get NP mark for the missed midterms or final or the missed questions, which will be used as 0 mark when calculating your final grade. No excuse will be accepted.
- UofR academic misconduct policies:  
<https://www.uregina.ca/president/executive-team/ed-governance-univ-secretary/student-appeals/student-behaviour.html>.
- Complaint about marking of assignments and quizzes should be made within 3 working days after the marks are posted (excluding the post date).
- As intellectual property, any part of any course work, including but not limited to assignments and exams, is not permitted to be posted to any website such as Chegg. Any posting to such a website that intends for help on a course work is considered as academic misconduct. Please be aware that many websites could trace and provide the information related to a posting upon request from academic institutions. For example, check the copyright policy of Chegg:

## Copyright Policy

Last updated on **July 27, 2020**

Chegg respects the intellectual property rights of others and we expect users of our websites and services to do the same.

Chegg is designed to support learning, not replace it. Misuse of our platform by any user may have serious consequences, up to and including being banned from our sites or having an academic integrity investigation opened by the user's institution. To learn more about our academic integrity policy, please visit <https://www.chegg.com/honorcode>.

In keeping with the Digital Millennium Copyright Act, we will respond promptly to valid written notification of claimed infringement regarding content posted on Chegg sites. Please note that Chegg may forward the written notification, including the complainant's contact information, to the user who posted the content. It is also our policy to disable and/or terminate the accounts of users who repeatedly infringe the copyrights of others.