CS4780: Information Retrieval
Course Policy

Hongning Wang
CS@UVa
http://www.cs.virginia.edu/~hw5x/Course/IR2021-Spring/_site/

Instructor

- Hongning Wang
  - Graduated from University of Illinois at Urbana-Champaign in 2014

Instructor

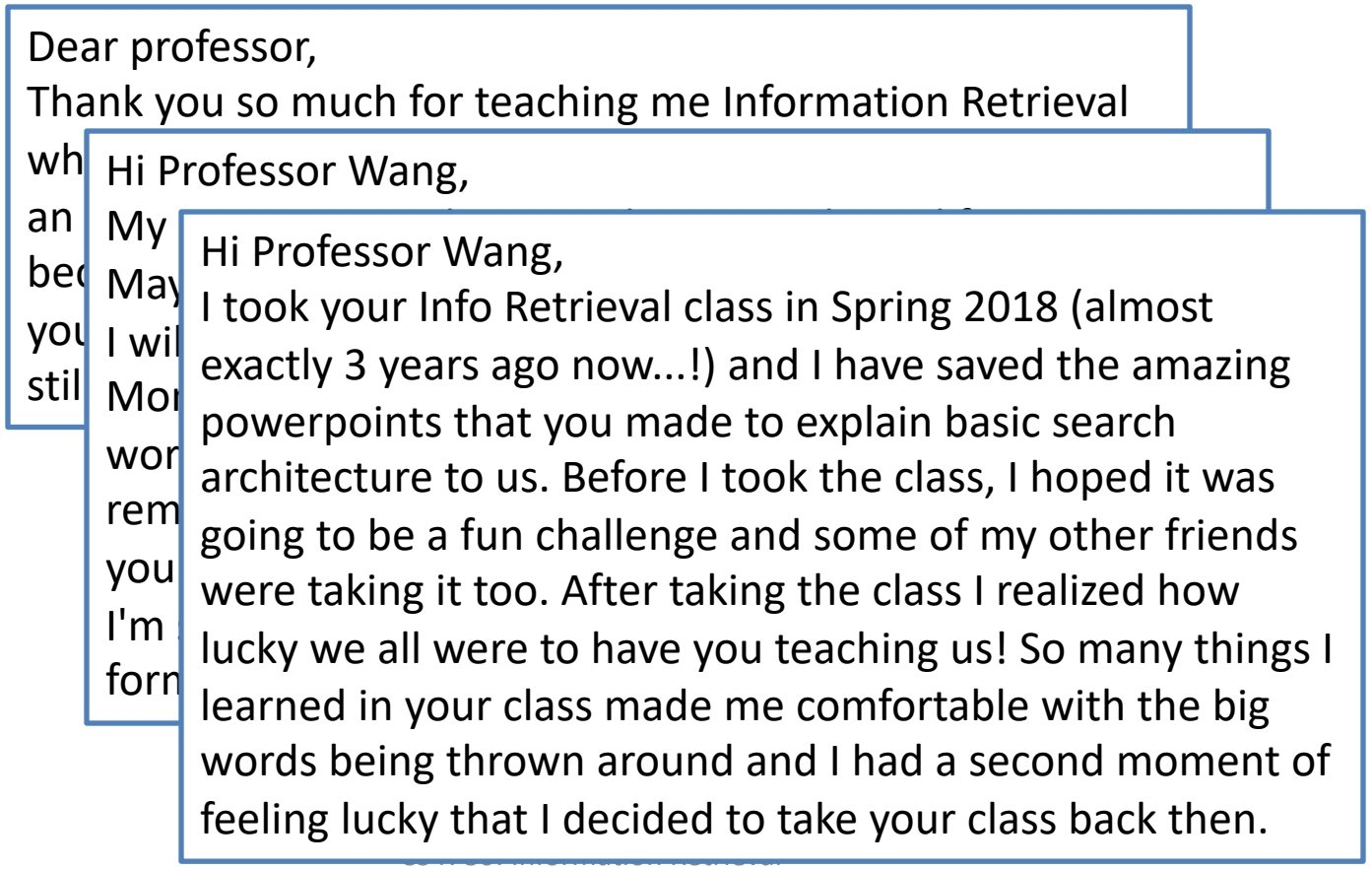
- Hongning Wang
  - Research area
    - Information retrieval
    - Data mining
    - Machine learning
  - Industry experience
    - Yahoo Labs
    - Microsoft Research
    - Snap
    - Google Research

Goal of this course

- Discuss fundamental problems in information retrieval
  - Building blocks of search engine systems
  - Wide coverage of important IR techniques
    - Personalized recommendation
    - Online advertising
- Get hands-on experience by developing practical systems/components
- Prepare students for doing cutting-edge research in information retrieval and related fields
  - Open the door to the amazing job opportunities in IT industry

Outcomes

- Letters from former students



Character of this course

- Discussion oriented
  - This is how great ideas are created!
  - You are encouraged to express your thoughts, confusions, and suggestions
  - Focusing on why, rather than how

Prerequisites

- Programming skills – Important!
  - Basic data structures: CS 2150 or equivalent
  - **Java** is required for machine problems
    - Most open source packages are written in Java
  - Any language you choose for the rest of this course
- Math background
  - Probability
    - Discrete/continuous distributions, expectation, moments
  - Linear algebra
    - Vector, matrix, dot product
  - Optimization
    - Gradient-based methods

Pop-up quiz

1. Let **a**=(1,2,3) and **b**=(2,3,-2), the inner product between **a** and **b** is

   (a) 0 (b) 1 (c) 2 (d) 3

2. Let A = $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$, what is A^{-1},

   (a) $\begin{pmatrix} -1 & -2 \\ -2 & -1 \end{pmatrix}$ (b) $\begin{pmatrix} -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{pmatrix}$ (c) $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ (d) $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$

Pop-up quiz

3. What is the expectation of random variables drawn from Gaussian distribution N(0, 1),

   (a) 0 (b) 0.5 (c) 1 (d) 2

4. Complexity of merge sort,

   (a) $O(n)$ (b) $O(n^2)$ (c) $O(\log n)$ (d) $O(n \log n)$

Pop-up quiz

1. Let **a**=(1,2,3) and **b**=(2,3,-2), the inner product between **a** and **b** is **(c)**

   (a) 0 (b) 1 (c) 2 (d) 3

2. Let $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$, what is A^{-1}, **(b)**

   (a) $\begin{pmatrix} -1 & -2 \\ -2 & -1 \end{pmatrix}$ (b) $\begin{pmatrix} -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{pmatrix}$ (c) $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ (d) $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$

Pop-up quiz

3. What is the expectation of random variables drawn from Gaussian distribution N(0, 1), **(a)**

   (a) 0 (b) 0.5 (c) 1 (d) 2

4. Time complexity of merge sort, ~~**(c)**~~ **(d)**

   (a) $O(n)$ (b) $O(n^2)$ (c) $O(\log n)$ (d) $O(n \log n)$

Structure of this course

- Six major topics will be covered by lectures
  - E.g., Search engine architecture, retrieval models, search evaluation, relevance feedback, and link analysis
- Latest development will be covered by paper reading assignments and presentations
  - E.g., mobile search, recommendation, personalization, online learning, you name it!

Grading policy

No curving will be applied!

- Reading assignments (10%)
  - Peer evaluation, after each chapter
- Homework (35%)
  - Machine problems (~3)
- Paper presentation (20%)
  - In class, performed in groups
- Course project (35%)
  - In the exam week

fairness will be guaranteed by the instructor

Reading assignments

- Read the instructor selected papers after each chapter
- Open-ended essay questions
- Peer evaluation on course forum

Paper presentation

- Choose to present the most recent works in the area of information retrieval
- Peer evaluation
- Choose from the instructor's selected papers, which are beyond our course content, so as to increase our topic coverage

Course project

- Topics
  - Implement algorithms in assigned research papers
  - Self-selected topics with permission from the instructor
- Team work
  - 3-4 students per group
- Evaluation
  - Two-page proposal (25%)
  - 15-minutes in-class presentation (40%)
  - Written report (35%)

In-class quiz

- After each chapter, the instructor will prepare a quiz to cover the most important concept in that chapter
- Its sole purpose is to help you review the learnt materials, and it is not part of our grading

Late policy

- Homework
  - Submit via Collab (no extension)
  - Late penalty: 15%, two weeks after the due date; 30%, afterwards
- Course project
  - Final report is due right after presentation (no extension)

Classroom participation

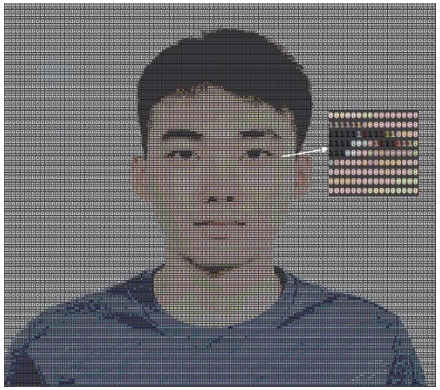
- HIGHLY APPRECIATED!
  - Helps me quickly remember your names
  - Reminds me what is still confusing
  - You can drive the lecture/discussion in this class!

Contact information

- Lecture
  - Instructor: Hongning Wang
  - Time: Tu/Th 2:00pm to 3:15pm
  - All via zoom, and recordings will be uploaded to collab right after
  - Office hours
    - Time: Tu/Th 3:30pm to 4:30pm
    - All via zoom, make appointments beforehand
    - Additional office hour can be requested by email

Contact information

- TA
  - Nan Wang (nw6a@virginia.edu)
  - Office hour
    - Time: Monday/Wednesday 10:30am to 11:30am
    - All via zoom, make appointments beforehand
    - Additional office hour can be request by email



Thank you!

QUESTIONS?