CS7263: Text Mining (Spring 2018)

Ferosh Jacob, fjacob1@kennesaw.edu

January 10, 2018

1 Course Goal

This course introduces the building blocks of information retrieval (IR) systems with necessary technical and mathematical background. On completion of the course, we should be able to: 1) design a new IR system for a given domain, 2) suggest methods to improve an existing IR system, and 2) evaluate an existing IR system. The course also aims to include discussions to understand state of the art web search systems.

2 Course Details

2.1 Textbook

Introduction to Information Retrieval by Christopher D. Manning Prabhakar Raghavan, and Hinrich Schütze (Free online version available here)

2.2 Meeting place

Atrium Building 157

2.3 Office hours

5:30 - 6:30 pm Wednesday. Please let me know early.

2.4 Meeting time

6:30 pm - 9:15 pm (Wednesday), To stay in schedule, we should have two lectures every Wednesday with a small break, say 15 minutes.

3 Course Prerequisites

3.1 Mathematical background

The students are expected to have a good grasp of the basic linear algebra and probability concepts.

3.2 Programming

This is a graduate level computer science course. It is highly recommended that the students should be fluent in either Java or Python.

4 Course Requirements

4.1 Read textbook

Students are expected to **read** the assigned text book chapters **before** the lecture.

4.2 Slides

Before the lecture, I will upload the slides in D2L which will be used for the lecture. But reading the slides is no substitute for reading the text book. Please read the textbook.

5 Mini-Projects

The course will include two mini-projects. All the students will be working on one problem but the problem will be somewhat open-ended allowing students to be creative. Each student is expected to submit the source-code and summary document and present a small demo of the implementation. I will upload a separate document detailing the instructions for the mini-projects.

6 Exams

We will have two exams: 1) Mid-term exam and 2) Final exam. The exams will be focused on the theory so it is highly unlikely that you will have coding questions for these exams.

7 Online discussion

For this course, we have online and in-class students and I really like to encourage online discussions about IR systems in general. In D2L, we have forums and I would like to see students using the forums for sharing IR relevant materials or general questions regarding the course.

8 Late Submission and Cheating Policies

Projects and Exams should be completed independently and any code not written by the student should be commented properly. I am planning to use Moss for detecting plagiarism.

9 Final Grade

The final grade will be determined as follows

20 %	Mid-term exam
25 %	Project 1
25 %	Project 2
20 %	Final exam
10 %	Class participation

What Will Happen If I'm Accused of Cheating/Plagiarism?[1]

This is what I usually tell faculty about how to handle plagiarism and other forms of cheating. Members of the faculty should confront and report academic dishonesty. To ignore cheating and plagiarism is to undercut the central mission of the university to educate.

- Professor detects alleged academic misconduct . Professor contacts Department of Student Conduct and Academic Integrity (SCAI) for advice and information about student's prior record. Phone: 470-578-3403 or email: scai@kennesaw.edu

- Professor may conduct disciplinary conference with student by him/herself or with the help of a facilitator.

- Professor sets disciplinary conference with student (this may take place at SCAI conference room if coordinated with SCAI Department.

- Professor or facilitator explains the structure of the meeting.

- Professor reviews the section of the Student Code of Conduct the student has allegedly violated and explains the nature of the accusation, providing all available evidence.

- Accused student receives opportunity to explain the situation and provide any evidence relevant to the explanation.

- Professor should wait until the student is finished to ask clarifying questions, providing reciprocal courtesy for the student's silence during the initial charge explanation.

- If student denies misconduct, professor determines if the explanation is satisfactory, in which case charges may be dropped. Educational dialogue should still usually take place prior to adjourning meeting.

- If student accepts responsibility, professor moves to educational dialogue on academic misconduct prior to discussion of sanctions.

- Educational dialogue should address the specific allegations as well as broader issues of academic misconduct.

---

[1]KSU plagiarism policy, http://scai.kennesaw.edu/students/general-info/cheating.php

- Following the dialogue, the professor states the academic sanctions deemed appropriate to the offense and explains the decision, taking the student's cooperation in the conference into consideration as a factor in determining severity.

- If student accepts responsibility, the professor completes the electronic academic integrity form. A representative from SCAI then sends a confirmation email to the student's KSU email that states the charge and resolution. The student has 10 business days after this confirmation letter is sent to contact SCAI if he or she believes the case has not been resolved as stated.

- Professor explains purpose of centralized records.

- If student denies misconduct and professor remains unconvinced, the meeting is immediately concluded and the matter referred to the Department of Student Conduct and Academic Integrity for a formal hearing.