Course Syllabus for CECS 429/529:
Search Engine Technology
Fall 2017

Department of Computer Engineering and Computer Science,
California State University, Long Beach

Instructor: Neal Terrell **Email**: neal.terrell@csulb.edu **Office**: VEC 405
Course hours and location: TuTh 9:30-10:20am VEC 519, 10:30-11:45am ECS 405
Office hours: TuTh 9:00-9:30am

Course Description

Models, algorithms, and evaluation of the retrieval of information from a collection of documents. Document preprocessing. Indexing and searching. Retrieval evaluation. Search engines.

Pre-requisites: CECS 323 and CECS 328, MATH 380 or EE 380

Topics

A general outline of topics to be covered:

- Basics of information retrieval
- Processing text documents
- Inverted index construction and compression
- Boolean information retrieval
- Ranked information retrieval and the vector space model
- Large-scale search engine construction
- Evaluation of search engine quality
- Text classification and machine learning
- Web search

Course Materials

Required: *Introduction to Information Retrieval*, Manning, C. Cambridge Press, 2008.

Supplementary material outside of the class textbook will be presented and included in tests, learning activities and programming projects. Links to additional material will be posted on the course website.

Grading

Students enrolled in CECS 529 will be subject to additional requirements outlined below, but will follow the same grading criteria.

Components:

- Homework - 0*
- Project - 25
- Midterms (x2) - 25
- Final - 25

Class grades will be determined using the traditional scale:

- A: Grade >= 90%
- B: 80% <= Grade < 90%
- C: 70% <= Grade < 80%
- D: 60% <= Grade < 70%
- F: Grade < 60%

Rules:

Homework will be assigned but not collected nor graded.

- I will give you problems I have written, problems from the book, or from online resources. These problems are for you to practice the topics covered in class.
- I will also give small programming problems, usually with a starting point that I have written. These problems are to help you break the very large project (see below) into more manageable pieces, and get early feedback on your coding decisions.

You will work on a large **project** over the course of the semester:

- You will implement (in code) and experiment with the ideas and methods presented in this course.
- You will work in teams of 2-3 people for significant periods of time.
- The project will be split into three milestones each with a strict due date. You will receive separate grades for each milestone, which all together will comprise your overall Project grade.
- Each milestone will be turned in electronically and will be demonstrated to me in person and/or presented to the class, depending on the milestone.
- Students enrolled in CECS 529 will be required to turn in additional work with each milestone.
- I will give you a grading rubric so you know how your project will be assessed.
- I expect a great deal from my students in general, and you should expect to work **very hard** on this project to meet my **exacting standards**.

Exams are closed-notes and closed-book. Please use the bathroom BEFORE taking the exam. BATHROOM BREAKS DURING EXAMS ARE NOT ALLOWED. MAKE-UP EXAMS ARE ONLY PROVIDED WHEN THERE IS DOCUMENTED EVIDENCE OF ACCIDENT OR ILLNESS.

Late penalties: project milestones will be assessed a 20% late penalty for each fraction of a calendar day that they are late.

Computer Software

You may use the Java, C++, or C# programming languages when answering homework problems or writing your projects. Other languages may be used **with explicit persmission of the instructor**.

Accessibility

DISABLED STUDENT SERVICES is a student support program within the Student Services Division. Our mission is to assist students with disabilities as they secure their university degrees at California State University, Long Beach. We provide services to over 13,000 students each semester. Over 3,000 students with disabilities have graduated from C S U Long Beach with support from our program.

The Disabled Student Services office is located on the 2nd floor in Brotman Hall, room 270.

It is your responsibility to notify the instructor in advance of any need for special accommodation due to a university verified disability.

Attendance and Drops

Attendance is not required, but all material presented during lecture or lab is fair game for exam questions. I will not redo a lecture for people who missed it the first time. If you miss a class day, it is your responsibility to obtain notes from someone who attended. I do not give "pop" quizzes.

Absences will not excuse you from turning in lab or project assignments on time. You are given plenty of time to complete these assignments, so pace yourself and plan to finish them early in case an emergency causes you to miss a day of class.

I will honor drop requests where permitted by university policy.

Academic Honesty

All assignments in this class are designated as *individual work only*. You may discuss ideas with others, but you may not share code, algorithms, or solutions with *any* individual on *any* of the class assignments. **Anything with your name on it must be written by you.** If you cannot complete an assignment on your own, the correct approach is to ask *me* for help during lab or office hours.

All instances of plagiarism or cheating, no matter how slight, will result in a 0 grade for the relevant assignment, and your final letter grade in the course will be reduced by one full grade.