Big Data Programming   CS {416, 616, 716}
Fall 2017   F 5:00 - 7:30 PM   CH 430

Instructor: Jeremy Blackburn (blackburn@uab.edu)

Office Hours: CH 141, MWF 2:00 - 3:00 PM and by appointment.

Textbook: NONE. We will be using just slides and academic papers.

Prerequisites: CS 303 and (MA 125 or MA 225)

Course Description:

This course is an introduction to working with Big Data. Big Data is a bit of a misnomer, but we are talking about "large" quantities of data, from a wide variety of sources, that tend to be continually updated, and from which we can extract *meaningful* knowledge. There are a variety of challenges associated with Big Data, from infrastructure, to programming paradigms, to analysis techniques. Even acquiring data is often somewhat of a challenge.

The goal of this class is thus to provide you with a toolbox from which you can build solutions related to Big Data. You will learn the basics of processing platforms and paradigms, storage solutions, computational methods, analysis techniques, and some programming languages and packages to help you work with Big Data.

It is important to note that in most cases the product of Big Data is *knowledge*; not just a program that does something. Therefore a primary goal of this class is to produce high quality reports on the work you do. In fact, these reports will make up the bulk of your grade.

Course Outline:

NB: This is the *intended* topics we will cover, but is subject to change depending on how class dynamics work out.

Note: 1 lecture does not necessarily correspond to 1 class period.

Assignments:

There are currently three medium sized projects and one final project planned, however, depending on how things go, this is somewhat subject to change.

Each project will be delivered as a report, in two column format. The three medium sized projects should result in approximately 2-4 page reports, while the final project should be 6-8 pages. Along with your reports, you will also submit source code where relevant. **We reserve the right to use plagiarism detection tools on both the submitted reports and submitted source code.**

Project descriptions will be made available as the class progresses along with specific formatting instructions. Graduate students will have a few extra things to include in their reports.

The final project will be driven by *you*. You will propose a topic, and using the things you learned in class, produce a report around that topic. You will also present your findings to the class (about 10-15 minutes).

In addition to the projects, graduate students will be required to submit written "reviews" of the research papers we cover in class. These reviews should provide evidence that you have put effort into understanding the research efforts of the larger community and how they provide context to this course.

Undergraduate students will not be required to submit these reviews, but will be awarded extra credit (up to 5% total, at our discretion) if they they do.

Grading: A Description of the grading process.
Undergraduate:

- Project 1 ....20%

- Project 2 ....20%

- Project 3 ....20%

- Final project 35%

- Class participation 5%

- Reviews (extra credit) ....up to 5%

Graduate:

- Project 1 ....20%

- Project 2 ....20%

- Project 3 ....20%

- Final project 35%

- Reviews 5%

Grade Policy:

Your final letter grade will be determined as follows:

$$90 \leq \text{A} \leq 100$$
$$80 \leq \text{B} < 90$$
$$70 \leq \text{C} < 80$$
$$55 \leq \text{D} < 70$$
$$0 \leq \text{F} < 55$$

Academic Honesty:

As a University of Alabama at Birmingham student, you have agreed to abide by the Universitys academic honor code. All academic work must meet the standards described in the UAB Academic Honor Code found at: `https://www.uab.edu/students/one-stop/policies/academic-honor-code` and also provided as a hard copy for you on the first day of class. Ignorance of the Academic Honor Code is *not* an excuse for violation. Questions related to course assignments and the Academic Honor Code should be directed to the instructor.

NB: A significant portion of my work has focused on cheating behavior. During my graduate studies I had a reputation for catching cheaters.

DO NOT RISK CHEATING. IT IS NOT WORTH YOUR REPUTATION AND ACADEMIC CAREER!

Extra Help:

Do not hesitate to come to my office during office hours or by appointment to discuss a project or any aspect of the course.

Attendance Policy:

I am *not* going to take attendance (except for first day). However, there *is* class participation that goes into your final grade, and there is very little chance that you will be given much leeway with respect to grading if I do not recognize you.

This class meets only once a week, so missing even a single class is a big deal.