Syllabus

Instructor: Jeff M. Phillips. | 3442 MEB | http://www.cs.utah.edu/~jeffp

Class Meetings: Mondays and Wednesdays, 3:00pm – 4:20pm, WEB 2230.

Course Web Page: http://www.cs.utah.edu/~jeffp/teaching/FoDA.html

This class will be an introduction to computational data analysis, focusing on the mathematical foundations. The goal will be to carefully develop and explore several core topics that form the backbone of modern data analysis topics, including Machine Learning, Data Mining, Artificial Intelligence, and Visualization. This will include some background in probability and linear algebra, and then various topics including Bayes Rule and connection to inference, linear regression and its polynomial and high dimensional extensions, principal component analysis and dimensionality reduction, as well as classification and clustering. We will also focus on modern PAC (probably approximately correct) and cross-validation models for algorithm evaluation.

Upon completion, students should have a basic understanding of the way modern data analysis goals are discussed, and grasp of the most standard approaches within the broad topics covered: regression, clustering, dimensionality reduction, and classification. Students who start with decent programming skills, should be able to implement one or more algorithms within each class.

The goal of this course is not to teach the state of the art, or the full array of techniques within these areas. Those goals are deferred to more advanced courses.

Book

The course text book will be Mathematical Foundations for Data Analysis by Jeff M. Phillips. A preprint is available here: www.mathfordata.com.

Getting Help

Take advantage of the instructor and TA office hours (posted on course web page). We will work hard to be accessible to students. Please send us email if you need to meet outside of office hours. Don't be shy if you don't understand something: come to office hours, send email, or speak up in class!

Students are encouraged to use a discussion group for additional questions outside of class and office hours. The class will rely on the Canvas discussion group. Feel free to post questions regarding any questions related to class: homeworks, schedule, material covered in class. Also feel free to answer questions, the instructors and TAs will also actively be answering questions. But, **do not post potential homework answers**. Such posts will be immediately removed, and not answered.

All important announcements will be made through the discussion group, there is otherwise no class mailing list.

Prerequisits

Some of the topics in this course are often taught as a "treat" at the end of a probability or linear algebra class, and then are often assumed knowledge in advanced data mining or machine learning classes. This class aims to fill these gap. This course will be a pre-requisite for advanced data analysis courses like Data Mining (CS 5140) and Machine Learning (CS5350).

We will begin setting prerequisites as CS 2100 (Discrete Math), CS 2420 (Algorithms & Data Structures), Math 2270 (Linear Algebra). And will set as a co-requisite CS 3130 (Probability and Statistics for

Engineers) or Math 3070 (Applied Statistics 1). These constraints are not formally in the system, but are recommended. Exceptions can be made for equivalent experience.

We will show and discuss code examples in python. No prior knowledge is expected.

Grading

There will be one final exam with 20% of the grade. Homeworks and quizzes will be worth 80% of the grade. There will be 5 homeworks and 4 quizzes – the lowest one (either one homework or one quiz) can be dropped. So each counted homework/quiz will be worth 10% of the grade.

The homeworks will usually consist of an analytical problems set, and sometimes light programming exercises in python. When python will be used, we typically will work through examples in class first.

Letter Grade Mapping: I will plan to map numerical grades to letter grades at the standard scale:

- 90-100 : A- to A
- 80-90 : B- to B+
- 70-80 : C- to C+
- 60-70 : D- to D+
- below 60 : E

The G- to G to G+ breakdown (for grade G = {A,B,C,D}) will probably align along:

- N0 to N3 : G-
- N3 - N7 : G
- N7 - N9.99 : G+

but I will reserve the right to shift this slightly.

Late Policy

To get full credit for an assignment, it must be turned in through Canvas by the start of class, specifically 2:45pm. Once the 2:45pm deadline is missed, those turned in late will lose 10%. Every subsequent 24 hours until it is turned another 10% is deducted. Assignments will not be accepted more than 48 hours late, and will be given a 0.

Assignments will be posted far enough ahead of time that I will not be able to make exceptions if a student falls ill. The exception will be prolonged illness, or other extenuating circumstance, (usually) accompanied by a doctors note.

If you believe there is an error in grading (homeworks or quizzes), you may request a regrading within **one week** of receiving your grade. Requests must be made by email to instructor, explaining clearly why you think your solution is correct.

Collaboration Policy

For assignments, you may discuss answers with anyone, including problem approach, proofs, and code. But all students must write their own code, proofs, and write-ups.

For quizzes and the exams, you must work by yourself. Students talking with other students during the tests will get a 0 score.

School of Computing Cheating Policy

The School of Computing has instituted a two strikes and youre out cheating policy, meaning if you get caught cheating twice in any SoC classes, you will be unable to take any future SoC courses. https://handbook.cs.utah.edu/2019-2020/Academics/policies.php

   If a student is caught cheating on a homework or the project, they will receive a failing grade for the course. For a detailed description of the university policy on cheating, please see the University of Utah Student Code: http://www.regulations.utah.edu/academics/6-400.html.

   Cheating on tests is swiftly enforced. Talking without an instructor or TA present during a test counts as cheating, and will be enforced on the spot. It will result in a zero grade on the test – it will be confiscated and not graded.

Students with Disabilities

The University of Utah seeks to provide equal access to its programs, services, and activities for people with disabilities. If you need accommodations in this class, reasonable prior notice needs to be given to the Center for Disability Services, 162 Olpin Union Building, 581-5020 (V/TDD). CDS will work with you and the instructor to make arrangements for accommodations.

Safety and Addressing Sexual Misconduct

The University of Utah values the safety of all campus community members. Title IX makes it clear that violence and harassment based on sex and gender (which Includes sexual orientation and gender identity/expression) is a civil rights offense subject to the same kinds of accountability and the same kinds of support applied to offenses against other protected categories such as race, national origin, color, religion, age, status as a person with a disability, veterans status or genetic in-formation. If you or someone you know has been harassed or assaulted, you are encouraged to come speak to the School of Computing Advisors and/or to the Title IX Coordinator in the Office of Equal Opportunity and Affirmative Action, 135 Park Building, 801-581-8365, or the Office of the Dean of Students, 270 Union Building, 801-581-7066. For support and confidential consultation, contact the Center for Student Wellness, 426 SSB, 801-581-7776.

   To report suspicious activity or to request a courtesy escort, call campus police at 801-585-COPS (801-585-2677). You will receive important emergency alerts and safety messages regarding campus safety via text message. For more information regarding safety and to view available training resources, including helpful videos, visit safeu.utah.edu. To report to the police, contact the Department of Public Safety, 801-585-2677(COPS). More information is available at https://safeu.utah.edu.

Latex

I highly highly recommend using LaTex for writing up homeworks. It is something that everyone should know for research and writing scientific documents. This linked directory (http://www.cs.utah.edu/~jeffp/teaching/latex/) contains a sample .tex file, as well as what its .pdf compiled outcome looks like. It also has a figure .pdf to show how to include figures.