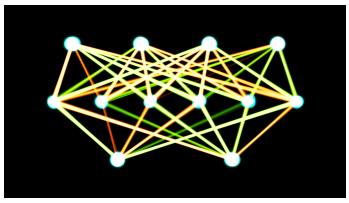
Fall 9-1-2021

CSCI 547.01: Machine Learning

Douglas John Brinkerhoff
University of Montana, Missoula, douglas1.brinkerhoff@umontana.edu

CSCI447/547
Machine Learning
TR, 9:30AM-10:50AM, SS 362



Instructor: Doug Brinkerhoff
E-mail: douglas1.brinkerhoff@umontana.edu
Office: SS403
Office Hours: T, 11:00AM-1:00PM, R, 1:00PM-3:00PM
(E-mail for an appointment, or my door is always open when I'm in.)
TA: Daithi Martin
E-mail: daithi.martin@umontana.edu
Office Hours: T, 1:00PM-3:00PM, R, 11:00AM-1:00PM

Course Description: As a society we have reached a point where the amount of information available to us exceeds our capacity to analyze it without the assistance of the very computers that have made the collection of such vast sums of data possible. Furthermore, this presents an exciting opportunity to understand our own cognition by programming machines to perform similar tasks. In this course we will explore the techniques required to turn data into predictive models of varying complexity, from simple (and foundational) Bayesian methods to neural networks, with many other methods in between. We will apply these methods to data drawn from biology, geoscience, natural language processing, and computer vision.

Course Objectives: At the completion of this course, the successful student will be able to:

1. Understand the language of machine learning in the context of contemporary data analysis.
2. Understand fundamental principles such as Bayesian updating, overfitting, and uncertainty.
3. Select models appropriate for the structure and subject of problems being considered.
4. Implement a bestiary of machine learning algorithms, both from scratch and with the assistance of high-quality libraries like scikit-learn and pytorch.

Class Organization: This course is flipped. You'll need to watch a video lecture, do the associated readings, and take a quiz prior to each class session. Class time will then be used to work collaboratively in groups to solve programming problems. The course will be taught using the programming language python, which is the de facto standard in the world of machine learning.

Course Webpages: The moodle page will be the central organizational repository, and all course materials will be linked from there, including video lectures, quizzes, and assignments. Assignments and my notes will typically be available as github repositories, and you should feel free to clone these repositories.

Student Evaluation: Students will be evaluated in several ways. Students will complete a quiz based on the course reading and video lecture before each class period. Students will also be graded on in-class assignments done in groups. There will also be 5 assignments that will generally *not* require programming, but rather will be oriented towards assessing your understanding of the conceptual aspects of the material. Grad students will have the enviable experience of answering an extra advanced problem on each assignment.

Graduate student will be required to propose and execute a project, which may be the implementation of a machine learning algorithm not covered in class, the application of machine learning to a non-trivial dataset, or an in depth report on a contemporary or classic academic paper on machine learning. For undergraduates the relative proportion of items contributing to final grades is 30% quizzes, 30% in-class work, 40% assignments. For graduate students the proportions are 20% quizzes, 20% in-class work, 30% assignments, and 30% project.

Computers, Software, and Online Material: If possible, bring a laptop to class. A tentative list of the software that we'll be using is as follows:

1. Python 3
2. Numpy/Scipy/Matplotlib: http://www.scipy.org/install.html
3. Jupyter: http://jupyter.org/install
4. scikit-learn: http://scikit-learn.org/stable/install.html

Prerequisite(s): Officially, CSCI232: Data Structures and Algorithms. In reality, this course requires a commitment to making up any knowledge gaps that the student might have with respect to the course material. Because of the nature of the subject, ML borrows heavily from topics in calculus, statistics, discrete math, and programming. It is unlikely that anyone is going to be comfortable with the course material all the time. Don't get too bent out of shape about it.

Text(s):

1. *Machine Learning and Pattern Recognition*, Christopher Bishop, freely available at `https://www.microsoft.com/en-us/research/uploads/prod/2006/01/Bishop-Pattern-Recognition-and-Mach pdf`

Late Assignments: I will not accept late assignments unless an extension was agreed upon well in advance of the due date or in extenuating circumstances to be determined at my discretion.

Academic Integrity: All students must practice academic honesty. Academic misconduct is subject to an academic penalty by the course instructor and/or a disciplinary sanction by the University. All students need to be familiar with the Student Conduct Code. I will follow the guidelines given there. In cases of academic dishonesy, I will seek out the maximum allowable penalty. Polemic: Look, this is a 400/500 level class, and if you're reading this you're probably looking to have a career in CS or a related field. When you're at a job interview, don't be sitting there regretting that you didn't learn anything in Machine Learning because you were cheating the whole time. Nobody wants that.

Disabilities: Students with disabilities may request reasonable modifications by contacting me. The University of Montana assures equal access to instruction through collaboration between students with disabilities, instructors, and Disability Services for Students. Reasonable means the University permits no fundamental alterations of academic standards or retroactive modifications.