Computer Science and Engineering 417T: Introduction to Machine Learning

Washington University in St. Louis, Spring 2020

Instructors:

Chien-Ju Ho chienju.ho at wustl dot edu

Class times:

Tuesdays and Thursdays 11:30am-12:50pm Wrighton / 300

1 Course Description

1.1 Overview

This course is an introduction to machine learning, focusing on supervised learning. We will cover the mathematical foundations of learning as well as a number of important techniques for classification and regression, including linear and logistic regression, neural networks, nearest neighbor techniques, kernel methods, decision trees, and ensemble methods. Note that the material in this course is a prerequisite for CSE 517A, the graduate level machine learning class. The overlap with CSE 511A (Artificial Intelligence) is minimal.

1.2 Prerequisites

CSE 247, ESE 326 (or Math 320), Math 233, and Math 309 (can be taken concurrently) or equivalents. If you do not have a solid background in calculus, probability, and computer science through a class in data structures and algorithms then you may have a hard time in this class. Matrix algebra will be used and is fundamental to modern machine learning, but it's OK to take that class concurrently.

1.3 Format

Class sessions will be lectures. There will be two in-class exams, one in early March (tentatively March 3) and on the last day of the class, April 23. There will be no separate final exam. There will be 5-7 homework assignments that will involve a mix of programming/computational exercises and pencil-and-paper problems. Homeworks will all be submitted online using Gradescope.

1.4 Textbooks

Most of the time we will not post lecture notes. Instead we will give references to the parts of the textbooks that correspond to the material covered in class on a given day. We may also post additional required or recommended reading to the website.

We will use the following textbooks:

1. *Learning From Data*, Y. Abu-Mostafa, M. Magdon-Ismail, and H-T Lin.
   `http://amlbook.com/`
2. *Computer Age Statistical Inference: Algorithms, Evidence, and Data Science*, B. Efron and T. Hastie.
   `https://web.stanford.edu/~hastie/CASI/`

1.5 Preliminary List of Topics

This preliminary list of topics may change based on time constraints and the progress of the class.

1. What is machine learning? Types of learning.
2. Generalization in finite and infinite hypothesis spaces. Training versus testing, model complexity, the VC bound, the bias-variance tradeoff.
3. Linear models: the perceptron, regression, logistic regression.
4. Nonlinear transformations of data.
5. The problem of overfitting. Regularization and validation as ways of preventing overfitting.
6. Modern supervised learning techniques, including decision trees, neural networks, nearest-neighbor methods, support vector machines, boosting, and random forests.

2 Policies

2.1 Announcements and Course Website

The main course website is at `http://chienjuho.com/courses/cse417t`. All announcements related to the class will be made either in lecture or on the website. **We will assume that any announcement made on the website is known to everyone in class within 24 hours of it being posted.** It is important to check the website regularly! You are responsible for all announcements made in lecture or on the website.

We will use Piazza for all questions and discussions related to the class. Please post questions on Piazza – they will reach the instructors and all the TAs, and you will get a quicker response. Individual emails about class issues will typically be met with a response saying "Please post your question to Piazza (anonymously if you so desire)." A link to the Piazza site will be on the main course website.

2.2 Assessment and Course Grade

We will first calculated your raw score using the following weights and determine your overall course score on a curve.

1. Homework assignments: 50%
2. Each exam: 25%

To make sure you are not penalized too much by a single under-performing homework assignment, your worst homework grades will be discounted by 50%.

If you would like to appeal your grade on any work, you may do so within 7 days of the work being handed back or the grade being received. In order to appeal the grade, please provide a detailed statement explaining why you believe the assigned grade is incorrect. We will regrade the entire piece of work, and your grade may go up or down, or it may stay the same.

2.3 Assignments and Late Days

Assignments will typically be due **by 11:59PM on the due date**. Each student will be given a budget of five *late days* that they can use to turn in homeworks late. A late-day can be used without explanation to extend a homework submission deadline by 24 hours, but **no more than two late-days can be used on any one homework**.

Any part of a late day that you use counts as a full late day. For example, if you do not submit your homework until 11PM if it is due at 10PM, that counts as a full late day. If you submit 26 hours late, you will have used two late days. You are responsible for keeping track of your usage of late days. **Use your late days wisely, if at all.** This late-day policy is intended to cover unanticipated things like minor sickness, exams in other classes, etc. so that you do not have to ask for extensions. Once you have used up your budget of late days you will not be allowed to turn in homeworks late for any reason other than true medical or family emergencies.

2.4 Collaboration and Academic Integrity

In this class, you are allowed to collaborate on assignments to the following extent. You are welcome to discuss problems with each other and to take your own notes during these discussions. However, you must write up solutions on your own. **You must write, in your submission, the names of students you discussed each problem with, and any external sources you used in a significant manner in solving the problem. Lack of citation of a source is a serious violation of this policy.** You may not use any solution keys or guides, however obtained, from previous versions of the class, similar classes, or the textbooks. You may not give or receive help from other students in the class on exams.

Submitting an assignment or exam that is in violation of this policy will automatically lead to receiving no credit for the assignment and a reduction of at least one grade modifier (e.g. from B to B-) beyond that in the overall course grade. However, depending on the circumstances, it could also lead to harsher penalties, for example, a failing grade in the class and initiation of the school's formal academic integrity review process. If you have any questions about the level of collaboration permitted, or any other aspect of this policy, please speak with the instructors about it before handing in the assignment!

2.5 Other accommodations and resources

Accommodations based upon sexual assault The University is committed to offering reasonable academic accommodations to students who are victims of sexual assault. Students are eligible for accommodation regardless of whether they seek criminal or disciplinary action. Depending on the specific nature of the allegation, such measures may include but are not limited to: implementation of a no-contact order, course/classroom assignment changes, and other academic support services and accommodations. If you need to request such accommodations, please direct your request to Kim Webb (kim_webb@wustl.edu), Director of the Relationship and Sexual Violence Prevention Center. Ms. Webb is a confidential resource; however, requests for accommodations will be shared with the appropriate University administration and faculty. The University will maintain as confidential any accommodations or protective measures provided to an individual student so long as it does not impair the ability to provide such measures.

If a student comes to us to discuss or disclose an instance of sexual assault, sex discrimination, sexual harassment, dating violence, domestic violence or stalking, or if we otherwise observe or become aware of such an allegation, we will keep the information as private as we can, but as faculty members of Washington University, we are required to immediately report it to our Department Chairs or Deans or directly to Ms. Jessica Kennedy, the Universitys Title IX Coordinator. If you would like to speak with the Title IX Coordinator directly, Ms. Kennedy can be reached at (314) 935-3118, jwkennedy@wustl.edu, or by visiting her office in the Womens Building. Additionally, you can report incidents or complaints to Tamara King, Associate Dean for Students and Director of Student Conduct, or by contacting WUPD at (314) 935-5555 or your local law enforcement agency.

You can also speak confidentially and learn more about available resources at the Relationship and Sexual Violence Prevention Center by calling (314) 935-8761 or visiting the 4th floor of Seigle Hall.

Bias Reporting The University has a process through which students, faculty, staff and community members who have experienced or witnessed incidents of bias, prejudice or discrimination against a student can report their experiences to the Universitys Bias Report and Support System (BRSS) team. See: http://brss.wustl.edu

Mental Health Mental Health Services professional staff members work with students to resolve personal and interpersonal difficulties, many of which can affect the academic experience. These include conflicts with or worry about friends or family, concerns about eating or drinking patterns, and feelings of anxiety and depression. See: http://shs.wustl.edu/MentalHealth