ECE 50024 Machine Learning

Course Information
Online Session

Lecture Hours: Asynchronous

Instructor: Professor Stanley Chan
Email: ece595chan@gmail.com
Office Hours: By email appointment.

Teaching Assistants:
Kent Gauen: ece595chan@gmail.com

Course Website: See Brightspace.

Course Objectives:

The objective of this course is that by the end of the semester, students will be able to

- Apply basic linear algebra, probability, and optimization tools to solve machine learning problems.
- Understand the principles of supervised learning methodologies, and can comment on their advantages and limitations.
- Explain the trade-offs in model complexity, sample complexity, bias, variance, and generalization error in the learning theory.
- Implement, debug, and execute basic machine learning algorithms on computers.

Catalog Description: Machine Learning is a dual-level (500-level) course customized for science and engineering students who are seeking to develop a solid mathematical foundation of the subject. The course focuses on the fundamental principles of machine learning instead of a scattered set of algorithmic tools. Students completing the course will be able to formulate the practical machine learning problems into mathematical frameworks, implementing algorithms to execute the statistical inference tasks, and analyzing the performance and limitations of the algorithms. The course has four parts: (1) linear regression, which covers regression models, outliers, ridge regularization, LASSO regularization, convex optimization, gradient descent algorithms, and stochastic algorithms; (2) classification, which covers separability, Bayesian classifiers, ROC curves, precision-recall curves, logistic regression, kernel methods; (3) learning theory, which covers probability inequality, the probably approximately accurate framework, generalization bound, model complexity, sample complexity, VC dimension, bias, variance, overfitting, and validation; (4) advanced topics of the state-of-the-arts, for example deep neural networks, generative models, and adversarial robustness. The course emphasizes the co-development of theory and programming. Students will have hands-on experience implementing machine learning algorithms in Python.

© 2023 Stanley Chan. All Rights Reserved.

Required textbook:

- *Introduction to Probability for Data Science*, by Stanley Chan, Michigan Publishing, 2021.
- *Learning from Data*, by Abu-Mostafa, Magdon-Ismail and Lin, AMLBook, 2012.

Recommended textbook:

- *Pattern Classification*, by Duda, Hart and Stork, Wiley-Interscience; 2 edition, 2000.
- *Elements of Statistical Learning*, by Hastie, Tibshirani and Friedman, Springer, 2 edition, 2009.
- *Pattern Recognition and Machine Learning*, by Bishop, Springer, 2006.

Pre-requisites: ECE20875 Python programming, and ECE302 Probabilistic Methods in ECE.

In addition to these minimum requirements, the following courses are highly recommended to students who are planning to take ECE 50024.

- **Linear Algebra** (as in the materials covered by G. Strang's Linear Algebra Textbook)

  A good course at Purdue is MA 511 Linear Algebra
- **Optimization** (as in chapter 1 - chapter 4 of S. Boyd's Convex Optimization)

  A good course at Purdue is ECE 647 (Week 1 - Week 8)
- **Probability** (as in the materials covered by D. Bertseka's Intro to Probability Textbook)

  A good course at Purdue is ECE 302 Probability. Graduate probability such as ECE 600 is recommended but not required.

To help you determine if you have adequate pre-requisites, we encourage you to try homework 0 posted in the homework section. If the problems are significantly beyond your comfort level, we suggest considering taking ECE 50024 at a later time.

Grades: All students will be graded by the following rubric. Everyone will be graded equally – no difference between Undergrad, MS, PhD. The final grade will be curved.

- Homework (20%). Homework are given approximately biweekly. Please submit your homework through gradescope. Late homework will not be accepted. You are encouraged to work in smalll groups, but you have to write / type your own solution. Worst homework will be dropped.
- We highly encourage you to type your solution using the LaTeX template provided in the course website, although we accept hand-written solutions. All programming answers should be typed.
- Quiz (30%). There will be 6 quizzes throughout the semester. The quizzes shall be taken on gradescope. Each quiz will be 30 minutes long. Worst quiz will be dropped.
- Project (50%). Please visit the course website for project instructions.

Academic Dishonesty:

© 2023 Stanley Chan. All Rights Reserved.

You are in college, not high school. We respect you as adults, and we expect you behave as adults. Therefore, we ask you to be honest and ethical in the course. In that respect, any action that might give a student unfair advantage on homework or exams will be considered dishonest. Examples include, but are not limited to:

- Sharing information during exam;
- Using forbidden material or device during exam;
- Viewing and/or working on an exam before or after the official time allowed;
- Requesting a re-grade of work that has been altered;
- Submitting work that is not your own. (You can discuss problems with your classmates. But you must write your own solution.)

All cases of academic dishonesty will be reported to the Office of Student Rights and Responsibilities, and will result in punishment. Possible punishments include, but are not limited to, a score of zero on work related to the cheating incident, a failing grade for the course, and, in severe cases, expulsion from the university.

Copyright of Course Material:
All ECE 50024 course material, including lecture, homework, project, solutions and exams are protected by copyright law. Without Prof Chan's permission, you are not allowed to distribute through any media including online sources. Below is an excerpt from `http://www.purdue.edu/studentregulations/student_conduct/misc.html`

... Students enrolled in, and authorized visitors to, Purdue University courses are permitted to take notes, which they may use for individual/group study or for other non-commercial purposes reasonably arising from enrollment in the course or the University generally. Notes taken in class are, however, generally considered to be "derivative works" of the instructor's presentations and materials, and they are thus subject to the instructor's copyright in such presentations and materials. No individual is permitted to sell or otherwise barter notes, either to other students or to any commercial concern, for a course without the express written permission of the course instructor...

FAQ:

1. I am a working adult. I find the math too difficult. Is there anything I can do?
    - If you feel that you are not ready for the course, please take it later.
    - The standard of the course is identical between online and residential sessions.
2. What are the pre-requisites of the course?
    - There is no official pre-requisite of the course (e.g., taking a prior course), although we expect students to have good background in linear algebra, optimization and probability.
    - Historically, undergraduate students and non-ECE PhD students have found this course difficult.
    - Please check out the information about pre-requisite to see if you are ready for the course.
3. What is the difference between ECE 50024 and other machine learning courses on campus?
    - We focus on general principles of learning.
    - Our goal is to provide an in-depth discussion of the subject, rather than superficially glancing through different topics.

© 2023 Stanley Chan. All Rights Reserved.

- We put significant emphasis on understanding the mathematics behind the algorithms.
- We have plenty of hands-on programming exercises.

4. What will I learn after taking ECE 50024?

   - You will know what a linear model is, such as Bayesian decision rule, perceptron algorithm, logistic regressoin, support vector machine, etc.
   - You will know how to understand a linear classifier from a geometric perspective.
   - You will know how to attack a classifier.
   - You will know the how much a machine learning algorithm can do, and what a machine learning algorithm cannot do.
   - You will know how to implement machine learning algorithms using Python and CVX.

5. I don't know how to type LaTeX. Can I submit word?

   - Yes for homework.
   - No for project.

6. Where can I get help for programming problems?

   - Please reach out to our teaching assistants.
   - If you feel that your Python programming skills are substantial below the minimum requirement, please consider taking a course on Python programming first.

© 2023 Stanley Chan. All Rights Reserved.