Fall 2020
Learning From Data

Instructor: Yang Li
TA: Weida Wang & Feng Zhao

Description

This introductory course gives an overview of many concepts, techniques, and algorithms in machine learning, beginning with topics such as logistic regression and SVM and ending up with more recent topics such as deep neural networks and reinforcement learning. The course will give the student the basic ideas and intuition behind modern machine learning methods as well as a bit more formal understanding of how, why, and when they work. The underlying theme in the course is statistical inference as it provides the foundation for most of the methods covered.

Intended Students

The course is geared towards students who are interested in understanding machine learning, and to carry out researches involving the applications of the machine learning problems. One of the objectives of the course is to understand the fundamental perspectives and develop solid connections between mathematical theory and learning systems.

Prerequisites

Basic concepts in calculus, probability theory, and linear algebra.

Problem Sets

There will be a total of 5 written and 4 programming problem sets, due roughly every 2-3 weeks. The content of the problem sets will vary from theoretical questions to more applied problems. You are encouraged to collaborate with other students while solving the problems but you will have to turn in your own solutions. Copying will not be tolerated. If you collaborate, you must indicate all of your collaborators. Each problem set will be graded by TAs.

Final Project

The final project for the course will involve using applied techniques on learning related applications or theoretical explorations of machine learning. The instructor will provide a list of suggested datasets for students to chose from, but students are encouraged to find their own dataset or topic, subject to the approval of the instructor.

The final project will be done in groups of two. Each group will submit a written report and optionally present in class[1].

Grading

Your overall grade will be determined roughly as follows:

ACTIVITIES	PERCENTAGES
Midterm	20 %
Final Project	30 %
Problem sets (written & programming)	50 %

Course Information

- Time: Friday 9:20-12:00am
- Location: Information Building 502
- Office Hour
  - Professor, Friday 2:00-4:00pm
  - TA Feng Zhao, Friday 20:00-22:00pm
  - TA Weida Wang, Friday 6:00-8:00pm

Reading Material

- Machine Learning Lecture Notes by Andrew Ng: `https://github.com/mxc19912008/Andrew-Ng-Machine-Learning-Notes`
- Pattern Recognition and Machine Learning (Book), by Christopher Bishop

Course Syllubus

Note: PA stands for "programming assignment"; WA stands for "written assignment".

Date	Topic	Homework release
9/18	Review Session (optional)	WA0 (don't need to hand in)
9/20	Introduction (make up for 9/18)	

[1] We may not have enough time for all groups to present during the final week. The final presentation method will be up to discussion later.

9/25	Supervised Learning I • Linear regression • Logistic regression	PA1
10/2	Chinese National Day	
10/9	Supervised Learning II • Generalized linear model • Model selection	WA1
10/16	Supervised Learning III • Generative model: GDA • Generative model: naive Bayesian model	PA2
10/23	Supervised Learning IV • Support vector machines	WA2
10/30	Supervised Learning V • Deep neural networks	PA3
11/6	Midterm	
11/13	Unsupervised Learning I • K-means clustering • Principal component analysis • Independent component analysis	WA3
11/20	Unsupervised Learning II • Canonical component analysis • Maximal HGR correlation	PA4
11/27	Unsupervised Learning III • Mixture Gaussian and EM algorithm • Spectral Clustering	WA4, Final Project
12/4	Machine Learning Theory I • Regularization • Empirical risk, VC dimension	
12/11	Machine Learning Theory II • Hypothesis testing	

12/18	Reinforcement Learning • Markov decision process • Value iteration and policy iteration • Q-Learning	WA5
12/25	Advanced Topic I • Transfer Learning	
1/1	Advanced Topic II • Semi-supervised learning	
1/8	Final Project Presentation I	
1/15	Final Project Presentation II	