Machine Learning CSCI 567, Spring 2023 Yan Liu

When: Tue 12:00-3:00pm
Where: GER 124
Office Hours: Wed 2-3pm (zoom link to be shared)
TAs: Pithayuth Charnsethikul, Liyu Chen, Samuel Griesemer, Chuizheng Meng

Overview

Objective: The major objective of this course is to introduce classical statistical machine learning methods, including but not limited to techniques for supervised and unsupervised learning problems. Particular focus will be laid on the conceptual understanding of these techniques, their applications and hands-on experience.
Prerequisites: (1) Undergraduate level training or coursework in linear algebra, multivariate calculus, basic probability and statistics; (2) Skills in programming with Python (self-studying numpy, scipy and scikit- learn and related packages is expected); (3) In addition, an undergraduate level course in Artificial Intelligence may be helpful but is not required.
Exercises: 4 written assignments + 4 programming assignments will be released as take-home exercises. These will not be graded and are for your own practice.
Grading: 3 Quizzes (20% each), homework (20%) - and a Final Project (20%).
Discussions: Attending the discussion sessions is required (they start from the second week). The discussion provides more detailed and in- depth exposition of the lecture materials as well as useful hints for take-home exercises.
Communication:
The main communication tool for this course is the D2L. (We will use D2L to distribute syllabus, lectures/videos/homework)

We will also use Piazza to host discussion forums and communicate with the instructors/TA team:

a) Any technical questions should be posted in the corresponding thread in the discussion forum. Your questions will be answered within 24 hours during the lifetime of the thread and 3-5 business days outside the lifetime of the thread.
b) Any non-technical questions can be posted as private messages to the instructor/TAs. We will respond within 24 hours.

We will email the link to Piazza soon.

Resources

Textbooks: There is no required textbook for this course, but the following textbooks are recommended readings:

- Machine Learning: A Probabilistic Perspective [MLaPP] by Kevin Murphy
- Elements of Statistical Learning [ESL] by Hastie, Tibshirani and Friedman
  Other useful resources:
  ML references:

- A First Encounter with Machine Learning by Max Welling
- Introduction to Machine Learning by Alex Smola and S.V.N. Vishwanathan

- An Introduction to Statistical Learning by Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani

- Bayesian reasoning and machine learning by David Barber
- Pattern Recognition and Machine Learning by Christopher Bishop

- A tutorial by Andrew Moore
  Math references:

- The Matrix Cookbook by Kaare Brandt Petersen and Michael Syskind Pedersen
- Calculus by Gilbert Strang
- Linear Algebra by Gilbert Strang
- Introduction to Probability and Statistics by Jeremy Orloff and Jonathan Bloom Learning Python:

• The official Python Tutorial

Schedule

*Schedule is subject to small adjustments. *Slides will be posted as the classes are held.

Syllabus and Materials

The following is a tentative schedule. The quiz timings are fixed, but the rest of the content will likely change as the course continues. We will also post lecture notes and assignments here.

Please refer to Ed Discussion for recommended readings.

Week	**Topics**	**Recommended Reading**
Week 1: 01/10	Course overview; Core ML concepts; nearest neighbor classification; typical steps of developing a ML system	
Week 2: 01/17	Self-assessment quiz on basic concepts that will be needed later this class; Linear regression; regression with nonlinear basis;	The exam will be graded, but it does not count towards the grade of the class.
Week 3: 01/24	Regularized regression Linear discriminant analysis	[MLaPP] 1.4.5, 7.1-7.3, 7.5.1, 7.5 7.5.4, 7.6, 1.4.7, 1.4.8 [ESL] 7.1, 7.2, 7.3, 7.10
Week 4: 01/31	Perceptron; logistic regression	[MLaPP] 4.2.1 - 4.2.5, 8.5.1-8.5.4 1.4.6, 8.1-8.3 [ESL] 4.1-4.2, 4.4
Week 5: 02/07	Multiclass classification; neural networks	[MLaPP] 16.5.1-16.5.6, 28 [ESL] 11.3-11.7
Week 6: 02/14	First quiz Convolutional neural net	
Week 7: 02/21	s; kernel methods	[MLaPP] 14.1, 14.2.1- 14.2.4, 14.4.1, 14.4.3 [ESL] 5.8, 6.3, 6.7
Week 8: 02/28	Lagrangian duality; SVM	[MLaPP] 14.5.2-14.5.4 [ESL] 12. 12.3
Week 9: 03/07	Second quiz Decision trees; boosting	[MLaPP] 16.4.1-16.4.5, 16.4.8, 16.4.9 [ESL] 16.3

Week 10: 03/14	No class -- Spring break	
Week 11: 03/21	Clustering; gaussian mixture models;	[MLaPP] 3.5, 11.1-11.3, 11.4.1-11.4.4, 11.5
Week 12: 03/28	gaussian mixture models; EM	[ESL] 6.6.3, 8.5, 14.3.1- 14.3.9,
Week 13: 04/04	Density estimation; generative models; naive bayes	[MLaPP] 17.1-17.4, 17.5.1-17.5.2
Week 14: 04/11	Hidden markov models; dimensionality reduction and visualization; PCA	[MLaPP] 10.1, 10.2.1- 10.2.3, 10.3-10.5, 12.2 [ESL] 14.5.1
Week 15: 04/18	Third quiz Multi-armed bandit	Materials will be distributed on D2L
Week 16: 04/25	Reinforcement learning Course project poster session	

Homework and Project

Homework	**Release Date**
PS1	01/24
PS2	02/19

PS3	03/06
PS4	04/03
Final Project	03/15

Problem sets: There are 4 problem sets, each containing one theory assignment and one programming task. These will be discussed in discussion sessions and solutions will be posted online later. These are not graded but we strongly encourage you to solve them in order to understand lecture material better and prepare for quizzes.

Quizzes

Quiz	Weight	Date and Time	Location
First Quiz	20%	02/14, 12-1:30pm	In-person
Second Quiz	20%	03/07, 12-1:30pm	In-person
Third Quiz	20%	04/18, 12-1:30pm	In-person

Logistics: Quizzes are individual effort, closed-book and closed-notes. You will be provided with papers. Any electronics including digital watches are not allowed. Any requirements for you to bring to the quiz will be announced well in advance. Students requiring alternate quiz arrangements must make such requests within the first two weeks of the semester, or as soon as possible after knowing of the conflict or requirement.

Academic Policies

Students with disabilities: Any student requesting academic accommodations based on a disability is required to register with Disability Services and Programs (DSP) each semester. A letter of verification for approved accommodations can be obtained from DSP. Please be sure the letter is delivered to the instructor or to a TA as early in the semester as possible.
Academic integrity: USC seeks to maintain an optimal learning environment. General principles of academic honesty include the concept of respect for the intellectual property of others, the expectation that individual work will be submitted unless otherwise allowed by an instructor, and the obligations both to protect one's own academic work from misuse by others as well as to avoid using another's work as one's own. All students are expected to understand and abide by these principles. Scampus, the Student Guidebook, contains the Student Conduct Code in Section 11.00. Students will be referred to the Office of Student Judicial Affairs and Community Standards for further review, should there be any suspicion of academic dishonesty.