CPTS 437: Introduction to Machine Learning (Spring 2018)

Course Overview:

Machine learning is the study of computer algorithms and models that learn automatically from data. It is a key area of artificial intelligence and has applications in many domains, including biology, social science, statistics and math etc. This introductory course covers key topics in machine learning, including linear models for regression and classification, generative models, support vector machines and kernel methods, neural networks and deep learning, decision trees, unsupervised learning and dimension reduction.

Course Information

- Lecture Time: Tuesday, Thursday, 2:50 PM - 4:05 PM
- Lecture Location: Sloan 233
- Required Textbook:
  - Learning from Data, by Yaser S. Abu-Mostafa, Malik Magdon-Ismail, and Hsuan-Tien Lin, Published by AMLBook
  - hard-copy containing Chapters 1-5 can be purchased at http://www.amazon.com/gp/product/1600490069
  - e-Chapters (6-9) can be downloaded from http://amlbook.com/
- Course announcements, lecture notes , homeworks will be posted on Blackboard.

Instructor

- Name: Shuiwang Ji
- Office: EME 123
- E-mail: sji@eecs.wsu.edu
- Telephone: (509) 335-4427
- Home page: `http://www.eecs.wsu.edu/~sji/`
- Office Hours: Tuesday, Thursday, 11:00 AM - 12:00 PM, or by appointment

Prerequisites

Required: CptS 223 or CptS 233 or CptS 215 (or equivalent). In addition, students are expected to have some familiarity with basic linear algebra (vectors, matrices, matrix-vector computations, vector and matrix norms, linear independence, matrix rank, singularity, positive definiteness, eigenvalues/eigenvectors, matrix decomposition, orthogonality), multivariate calculus (derivatives of univariate functions, derivatives of multivariate functions, chain rule, Taylor expansion),

and basic probability and statistics (discrete and continuous probability distributions, sum rule, product rule, marginal probability distributions, conditional probability distributions, joint probability distributions, independence and conditional independence, Bayes Theorem, variance and covariance, expectation).
Recommended: EE 221

Grading

- Homework (6): 35%
- Exam 1: 20%
- Exam 2: 20%
- Final (5/1/2018, 10:10 – 12:10 PM): 25%

Final letter grades will be assigned based on absolute percentage as follows:

Absolute percentage	[100, 93]	(93, 90]	(90, 86]	(86, 83]	(83, 80]	(80, 76]
Letter grade	A	A-	B+	B	B-	C+

Absolute percentage	(76, 73]	(73, 70]	(70, 66]	(66, 60]	(60, 0]
Letter grade	C	C-	D+	D	F

where [ ] denotes inclusion and ( ) denotes exclusion. The instructor reserves the right to move the thresholds down (but not up) based on the distribution of final percentages.

Week-to-Week Schedule (Subject to Change)

- Week 1: Introduction
- Week 1: Linear algebra review
- Week 2: Linear regression (homework 1 assigned)
- Week 3: Logistic regression (homework 1 due)
- Week 4: Probability review (homework 2 assigned)
- Week 5: Basic probability and naïve Bayes classifier (homework 2 due)
- Week 6: Generative versus discriminative models (Exam 1)
- Week 7: Generalization and overfitting (homework 3 assigned)
- Week 8: Support vector machines and kernel methods (homework 3 due)
- Week 9: A unified view of supervised learning
- Week 10: Neural networks (homework 4 assigned)
- Week 11: Deep learning (Exam 2) (homework 4 due)
- Week 12: Multi-class learning (homework 5 assigned)

- Week 13: Decision tree and random forests (homework 5 due)
- Week 14: Dimension reduction, principal component analysis (homework 6 assigned)
- Week 15: Unsupervised learning and K-means clustering, spectral clustering (homework 6 due)
- Final exam

Student Learning Outcomes and Assessment

Student learning outcomes include (1) understanding the foundation, major techniques, applications, and challenges of machine learning; (2) the ability to apply basic machine learning algorithms for solving real-world problems. The learning outcomes will be assessed based on a combination of homework assignments and exams.

Student Learning Outcomes	Course Topics/Dates	Evaluation
Understand linear regression algorithms	Weeks 1&2	Assignments, Exams
Understand linear classification algorithms	Weeks 3,4,5	Assignments, Exams
Understand generative/discriminative models	Week 6	Assignments, Exams
Understand generalization	Week 7	Assignments, Exams
Understand kernel methods	Week 8,9	Assignments, Exams
Understand neural networks and deep models	Weeks 10,11,12	Assignments, Exams
Understand decision trees	Week 13	Assignments, Exams
Understand unsupervised learning methods	Week 14	Assignments, Exams
Understand clustering methods	Week 15	Assignments, Exams

Homework

All homework must be done independently, or you will be penalized for plagiarism. The instructor and the TA will be carefully looking into your code.

Most homework contains a written component and a programming component. Therefore, most homework submission should include a report and code. Submission instructions will be provided on each homework assignments. The homework should be submitted BEFORE class on the due date. Late penalty is 15% point deduction per day for the first three days (including weekends), after which the submission will not be accepted. Exceptions/extensions can be given to students with valid excuse. Students need to provide evidence for their excuse and must notify the instructor before the original due date. A student may request up to 2 excused extensions.

Most homework requires Python programming. Data and skeleton code will be provided in Python format.

Exams

Students will be required to complete two in-class exams and one final exam. In-class exams will focus on topics taught since the last exam. The final exam will be comprehensive.

Class Participation

Students are required to attend all classes and actively participate in discussions. The instructor will use whiteboard extensively, thus class participation is required and essential.

Academic Integrity

WSU definitions and procedures for cases of academic dishonesty are given at `conduct.wsu.edu`. These procedures will be followed rigorously. All work submitted for grading is to be original. Material submitted that is not original must be cited as described in technical writing text books. No copying from other sources will be accepted. Students who violate this policy will receive a zero grade on the assignments in question and/or receive a failing grade for the course.

Students with Disabilities

Reasonable accommodations are available for students with documented disability. If you have a disability and may need accommodations to fully participate in this class, please visit the Access Center (Washington Building 217) to schedule an appointment with an Access Advisor. All accommodations MUST be approved through the Access Center. Additional information can be viewed at `http://drc.wsu.edu`.

Campus Safety

Classroom and campus safety are of paramount importance at Washington State University, and are the shared responsibility of the entire campus population. WSU urges students to follow the "Alert, Assess, Act" protocol for all types of emergencies and the "Run, Hide, Fight" response for an active shooter incident. Remain ALERT (through direct observation or emergency notification), ASSESS your specific situation, and ACT in the most appropriate way to assure your own safety (and the safety of others if you are able). The Campus Safety Plan, which can be found at `http://safetyplan.wsu.edu`, contains a comprehensive listing of university policies, procedures, statistics, and information relating to campus safety, emergency management, and the health and welfare of the campus community.