CSE512: Machine Learning

Spring 2021, Time: Tues & Thurs 9.45am – 11.05am, Zoom Online
Instructor: Minh Hoai Nguyen, New CS 243.
Office hours: Thursdays and Fridays 5.15-6.00pm. Location for in-person sessions: NCS 120.
TAs: TBA.

`Click here for Zoom link` You will need to sign in with your NetID Google Account to view.

> This page provides some information to students who are considering to enroll in the course. Below are tentative grading scheme and syllabus, which are subject to change. Once you have enrolled in the course, you should go to Blackboard for updated information and futher annoucements. The content of this page will not be updated two weeks after the course has begun.

This course covers fundamental machine learning concepts for intelligent systems that autonomously learn to perform a task and improve with experience, including problem formulations (e.g., selecting input features and outputs) and learning frameworks (e.g., supervised vs. unsupervised), standard models, methods, computational tools, algorithms and modern techniques, as well as methodologies to evaluate learning ability and to automatically select optimal models. Applications to areas such as computer vision (e.g., image classification, object detection) and medical data analysis (e.g., Covid-19 prediction) will motivate the coursework and material.

This course is intended for graduate students who already have good programming skills and adequate background knowledge in mathematics, including probability, statistics, and linear algebra. This course is offered by the Department of Computer Science, and students from the department will have priority in registering for this course. If there are still spaces, students from AMS and BMI will be given special permission to enroll. Students from other departments won't be able to enroll in this course due to resource constraint. For special permission to enroll, attend the few two weeks of lectures (using the Zoom link provided in the Google doc above), and I will tell you how to enroll in the course at the end of the second week if: 1) you are still interested in taking the course, and 2) spaces are still available.

Grading

There will be 5-6 homework assignments, two midterm exams, and several class quizzes.

- Homework assignments: 50%
  - Five to six homework assignments
- Exams and Quizzes: 50%
  - Class quizzes: 10%
  - Midterm 1: 20%
  - Midterm 2: 20%
- No final exam. No final project.

Weights are approximate and subject to change. You are expected to do homework assignments by yourselves. Even if you discuss them with your classmates, you should turn in your own code and write-up. You can have one sheet of paper with notes in the exams.

Tentative Syllabus

Week	Date	Topic	Assignment/Exam	In-person Check-in
1	2-Feb-21	Class cancelation due to snow storm		
	4-Feb-21	Introduction, MLE and MAP		
2	9-Feb-21	Linear Regression	HW1 out	
	11-Feb-21	Bayes Risk and Nearest Neighbor Classifier		
3	16-Feb-21	Naïve Bayes, Logistic Regression		
	18-Feb-21	Generative versus Descriminative Classifier		
4	23-Feb-21	Bias-Variance Tradeoff	HW1 due. HW2 out	
	25-Feb-21	Performance evaluation		
5	2-Mar-21	Regularization, Ridge regression		NCS120
	4-Mar-21	LASSO regression and sparsity		
6	9-Mar-21	Max-margin learner and primal SVM	HW2 due. HW3 out	
	11-Mar-21	No lecture - Midterm 1	Midterm 1	
7	16-Mar-21	Duality and Dual SVM		
	18-Mar-21	Kernel tricks and Kernel SVM		
8	23-Mar-21	Decision Trees	HW3 due. HW4 out	
	25-Mar-21	Ensemble learning - Bagging		
9	30-Mar-21	Ensemble learning - Boosting		
	1-Apr-21	Unsupervised learning, K-means and PCA		NCS120
10	6-Apr-21	GMM and EM algorithm	HW4 due. HW5 out	
	8-Apr-21	Hidden Markov Models		
11	13-Apr-21	Hidden Markov Models		
	15-Apr-21	Artificial Neural Networks		
12	20-Apr-21	Training Deep Networks	HW5 due. HW6 out	
	22-Apr-21	Convolutional Neural Network		
13	27-Apr-21	Convolutional Neural Network		
	29-Apr-21	No lecture - Midterm 2	Midterm 2	
14	4-May-21	Recurrent Neural Network	HW6 due	NCS120
	6-May-21	Final lecture - course review		

Students who are enrolled in the in-person/hybrid session of this course must attend at least two out of three in-person sessions, which are highlighted in the above table. Students who fail to do so will not receive a passing grade. In-person sessions are held during the regular class hours, in Room 120 of the New Computer Science building (NCS 120). You must physically be in NCS 120 during the in-person sessions, but you will be watching synchronous Zoom lectures on your own computers. Bring your own computers and headphones.

Textbooks

Textbooks are optional.

Books that I love and recommend:

- Tom Mitchell, Machine Learning, McGraw Hill, 1997.
  - Excellent book. Concise and well-explained. But it is quite old and does not cover recently developed concepts.
  - Some New Chapters are available
- Mohri, Rostamizadeh, Talwalkar, Foundations of Machine Learning, MIT Press, 2nd edition, 2018.
  - Excellent book. Well-explained. Deep and mathematically rigorous. Suitable for mathematically inclined students.
  - Free PDF is available
  - Use discount code MTSR20 for 30% discount from MIT press
- Sutton & Barto, Reinforcement Learning: An Introduction, MIT Press, 2nd edition, 2018.
  - Excellent book on reinforcement learning. Well-explained concepts with lots of examples. This provides a systematic view of multiple RL approaches.
  - For reinforcement learning, this is the book to read.

Other textbooks that may be more useful or accessible:

- Trevor Hastie, Robert Tibshirani, Jerome Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd Ed, Springer Series in Statistics, 2009.
- Christopher M. Bishop, Pattern Recognition and Machine Learning, Springer, 2011.
- Kevin Murphy, Probabilistic Machine Learning: An Introduction, MIT Press 2021

Student Accessibility Support Center Statement

If you have a physical, psychological, medical, or learning disability that may impact your course work, please contact the Student Accessibility Support Center, 128 ECC Building, (631) 632-6748, or via e-mail at: sasc@stonybrook.edu. They will determine with you what accommodations are necessary and appropriate. All information and documentation is confidential.

Academic Integrity Statement

Each student must pursue his or her academic goals honestly and be personally accountable for all submitted work. Representing another person's work as your own is always wrong. Faculty is required to report any suspected instances of academic dishonesty to the Academic Judiciary. Faculty in the Health Sciences Center (School of Health Technology & Management, Nursing, Social Welfare, Dental Medicine) and School of Medicine are required to follow their school-specific procedures. For more comprehensive information on academic integrity, including

categories of academic dishonesty please refer to the academic judiciary website at http://www.stonybrook.edu/commcms/academic_integrity/index.html

Critical Incident Management

Stony Brook University expects students to respect the rights, privileges, and property of other people. Faculty are required to report to the Office of University Community Standards any disruptive behavior that interrupts their ability to teach, compromises the safety of the learning environment, or inhibits students' ability to learn. Faculty in the HSC Schools and the School of Medicine are required to follow their school-specific procedures. Further information about most academic matters can be found in the Undergraduate Bulletin, the Undergraduate Class Schedule, and the Faculty-Employee Handbook.

Attendance and Covid-19

Students are expected to attend every class, report for examinations and submit major graded coursework as scheduled. If a student is unable to attend lecture(s), report for any exams or complete major graded coursework as scheduled due to extenuating circumstances, the student must contact the instructor as soon as possible. Students may be requested to provide documentation to support their absence and/or may be referred to the Student Support Team for assistance. Students will be provided reasonable accommodations for missed exams, assignments or projects due to significant illness, tragedy or other personal emergencies. In the instance of missed lectures or labs, the student is responsible for insert course specific information here (examples include: review posted slides, review recorded lectures, seek notes from a classmate or identified class note taker, write lab report based on sample data). Please note, all students must follow Stony Brook, local, state and Centers for Disease Control and Prevention (CDC) guidelines to reduce the risk of transmission of COVID. For questions or more information click here.