Syllabus

CS 688

Machine Learning

Instructor

Antonios Anastasopoulos (antonis [at] gmu [dot] edu)
Office Hours: ENGR 4412, Tue 2-3pm.

Teaching Assistant

Ted Chao (cchao8 [at] gmu [dot] edu)
Office Hours: Thursdays 10-12am, Blackboard sessions.

Meets

Thursday, 4:30 to 7:10 PM, Planetary Hall 206.
Safe Return to Campus: Students are expected to follow the university's Safe-Return-to-Campus Policy (including mask wearing, daily health check, etc.) for attending any classes. Please check out the policy before coming to the campus and the classroom. Note that students who choose not to abide by these expectations will be referred to the Office of Student Conduct for failure to comply.
If you are experiencing COVID-like symptoms, or if you have been in contact with a known case, please DO NOT put others in danger and DO NOT come to the class. Any classes missed due to COVID will be excused, and you'll be able to watch the recorded lecture at your convenience later.

Textbook

[no textbook is required but using either of the books below is highly recommended!]
C. M. Bishop Pattern Recognition and Machine Learning, Springer, 2006.
Additional books that you might find interesting:

- Machine Learning: A Probabilistic Perspective by Kevin P. Murphy
- Machine Learning by Tom Mitchell (link).
- A Course in Machine Learning by Hal Daumé III (link)
- Human-in-the-Loop Machine Learning by Robert (Munro) Monarch

Course Web Page

https://nlp.cs.gmu.edu/course/cs688-spring22/.
We will use **Blackboard** for course materials, Gradescope for assignments/grading, and **Piazza** for Q&A (sign up link: TBA).

Course Description

Machine learning studies computer algorithms for learning to do things. For example, we might be interested in learning to complete a task, or to make accurate predictions, or to navigate in an unexplored environment. The learning that is being done is always based on some sort of observations or data, such as examples (the most common case in this course), direct experience, or instruction. So in general, machine learning is about learning to do better in the future based on what was experienced in the past. The emphasis of machine learning is on automatic methods. In other words, the goal is to devise learning algorithms that do the learning automatically without human intervention or assistance.
The machine learning paradigm can be viewed as "programming by example." Often we have a specific task in mind, such as recognizing handwritten digits on an envelope to perform automated mail dispatching. But rather than program the computer with rules to solve the task directly, in machine learning, we seek methods by which the computer will come up with its own program based on examples that we provide.

The course covers key algorithms and theory at the core of machine learning. Particular emphasis will be given to the statistical learning aspects of the field. Topics include: decision theory, Bayesian theory, curse of dimensionality, linear and non-linear dimensionality reduction techniques, classification, clustering, neural networks, kernel methods, mixture models and EM, ensemble methods, deep learning.

Prerequisites

- CS 580 or CS 584 or permission of instructor.
- Students should be experienced with writing substantial programs in Python.
- Students must be familiar with basic probability and statistics concepts, linear algebra, optimization, and multivariate calculus.

Please contact the instructor if you have questions about the necessary background.

Class Format

Lectures by the instructor. Besides material from the textbook, topics not discussed in the book may also be covered. Research papers and handouts of material not covered in the book will be made available. Grading will be based on participation, homeworks, and a midterm exam. Homework assignments will be given and the solutions will be discussed in class in the following lectures. In order to learn the material and to do well on quizzes and the exam, students are required to work on the assignments. Graded work must be done on an individual basis, unless otherwise stated by the instructor. Any deviation from this policy will be considered a violation of the GMU Honor Code.

Classroom Specifics

I expect students to attend the class. I will supplement the textbook with extensive discussions and material. Students' active participation is very important to succeed in this course.
Some classes will include in-class exercises that will build on the material taught. These will be handouts (or will be small programming assignments) that I will ask you to complete in class. They will not be graded for their accuracy, but participating in the exercise and submitting something back will reflect your participation grade, along with short quizzes on the discussed material.

Grading

There will be a small midterm exam (no final exam). Your final grade will be dependent on:
Homeworks (60%): There will be 8-10 homework assignments, scatterred throughout the semester. Expect to have something due every week of the semester. Some of the homeworks will be theoretical (pen-and-pencil exercises), some of them will involve programming exercises. Details and materials will be published on Blackboard.

Participation (10%): See details above.

Midterm (30%): TBA.

Late Day Policy: Late homework submissions are only eligible for 90% of the points the first day (24-hour period) after the deadline, 80% the second, 70% the third, and 60% the fourth. That is, your score for that homework will be $min(90, s)$ if you submit one day late, $min(80, s)$ if you submit two days late, etc, where s is your raw score.

You receive 5 total grace days for use on any homework assignment except HW0. We will automatically keep a tally of these grace days for you; they will be applied greedily. No assignment will be accepted more than 4 days after the deadline. This has two important implications: (1) you may not use more than 4 graces days on any single assignment (2) you may not combine grace days with the late policy above to submit more than 4 days late.
All homework submissions are electronic. As such, lateness will be determined by the latest timestamp of any part of your submission.

Extensions In general, we do not grant extensions on assignments. However…

- Covid-19 is still here. This syllabus can't pre-solve all the situations that could arise.
- **please keep in communication** when things are getting rough - if you or people you take care of at home are ill and need all your attention; if you are facing evictions or problems that will impact your ability to participate in class and do the assignment work. We can strategize and think it through far better ahead of time than afterwards, for sure. We can affect the future far more effectively than the past!

Contested Grades All requests are due within a week of the grade becoming available on Blackboard. To do so, either schedule a meeting in person or send an email requesting further feedback and consideration. After that week, the window to contest a grade has closed other than recording errors. Contact the GTA about homework, and contact the professor about tests.

Readings

For some topic/class the instructor will provide a list of papers as suggested readings. For some classes later in the semester, one paper will be required reading and will be tested with a quiz (see above). Students should be able to understand the course content just by following the lecture along with the textbook and by doing the readings.

Tentative Schedule

Week	Date	Topic	Homework Due
1	1/27	**Introduction** • Overview • Course outline and syllabus • Formalizing the Learning problem • Limits of Learning • Probability Theory • Information Theory • Decision Trees **Lab:** • Background exercises	HW0: Background out. Due 2/2
2	2/3	**Classification** • Decision Trees (Part II) • k Nearest Neighbors • the Curse of Dimensionality • Perceptron • Feature Pruning and Normalization • Evaluating Model Performance • Bias/Variance Trade-off **Lab:** • kNN	HW1 Due 2/2 HW1: DTs and Overfitting out.
3	2/10	**Linear Models** • Linear Regression • Gradient Descent • Weight Regularization • Support Vector Machines • Fairness and Data Bias **Lab:** • Gradient Descent	HW2 Due 2/9 HW2: Linear Models for Classification out.

4	2/17	**Probabilistic Modeling** • Estimating Probabilities: MLE and MAP • Lagrange Multipliers • Generative vs Discriminative • Logistic Regression **Lab:** • SVMs	HW2 Due 2/16. HW3: Logistic Regression out.
5	2/24	**Probabilistic Modeling (II)** • Naive Bayes Models • Generative Stories • Joint and Conditional Models • Regularization via Priors **Kernel Methods** • Kernels and the Kernel Trick • Representer Theorem	HW3 Due 2/23. HW4: Kernels out.
6	3/3	**Neural Networks** • Gradient Descent • Backpropagation • Optimization Algorithms • Overfitting • Generalization Error • Early stopping **Lab:** • Basic NNs	HW4 due 3/2. HW5: Neural Networks out.
7	3/10	**Deep Learning** • Convolutional Neural Networks • Parameter Sharing • Recurrent NN • Applications and Limitations **Lab:** • Deep Learning with pyTorch	HW5 Due 3/12.
8	3/17	**NO CLASS (Spring Recess)**	
9	3/24	**Midterm Exam**	
10	3/31	**Unsupervised Learning** • K-means Clustering • Expectation-Maximixation • Mixture Modeling **Semi-supervised Learning**	HW6: EM out.
11	4/7	**Learning Paradigms** • PCA and Dimensionality Reduction **Lab:** • PCA	HW6 due 4/6.
12	4/14	**Structured Prediction** • Hidden Markov Models • Viterbi Algorithm • Conditional Random Fields **Lab:** • CRFs with crfsuite	
13	4/21	**Reinforcement Learning** • Markov Decision Processes • Exploration vs Exploitation **Lab:** • Reinforcement Learning	
14	4/28	**Human-in-the-Loop ML** • Active Learning • Uncertainty Sampling • Diversity Sampling • Sample Applications	
15	5/5	**Special Topics**	

Honor Code

The class enforces the GMU Honor Code, and the more specific honor code policy special to the Department of Computer Science. You will be expected to adhere to this code and policy.

Statement on Inclusion

I value the many perspectives that all of you bring to our class. I value each and every one of you. I want you to succeed, to feel comfortable, to be seen and heard. Please help cultivate the supportive, positive environment for everyone in class that we all deserve. We are allowed to be unique individuals. Be patient and kind with each other as we all work through another interesting semester in each other's company. Assume the best of each other and our intentions, and own up to the effects we have on others. I will do my best to pay attention to how you have arrived at this course, meet you where you are, and help you get the most out of our time together. If you work harder, I will meet your efforts with you. If you feel lost or unsupported, I will help you when I can and help you find other supports that go beyond me.

Note to Students

Take care of yourself! As a student, you may experience a range of challenges that can interfere with learning, such as strained relationships, increased anxiety, substance use, global pandemics, feeling down, difficulty concentrating and/or lack of motivation. All of us benefit from support during times of struggle. There are many helpful resources available on campus and an important part of having a healthy life is learning how to ask for help. Asking for support sooner rather than later is almost always helpful. GMU services are available, and treatment does work. You can learn more about confidential mental health services available on campus at: https://caps.gmu.edu/. Support is always available (24/7) from Counseling and Psychological Services: 703-527-4077.

Learning Disabilities

If you have a documented learning disability or other condition which may affect academic performance, make sure this documentation is on file with the Office of Disability Services and come talk to me about accommodations. I will work with you to ensure that accommodations are provided as appropriate. If you suspect that you may have a disability and would benefit from accommodations but are not yet registered with the Office of Disability Services, I encourage you to contact them at ods@gmu.edu.

NEXT
Homework

Last updated on Oct 5, 2021