CM146: Introduction to Machine Learning

Instructor: Kai-Wei Chang

Course description

Machine Learning encompasses the study of algorithms that learn from data. It has been a key component in a number of problem domains including computer vision, natural language processing, computational biology, and robotics. This class will introduce the fundamental concepts and algorithms in machine learning (supervised as well as unsupervised learning) as well as best practices in applying machine learning to practical problems. The class consists of lectures, problem sets that contain mathematical and programming exercises, midterm, quizzes, and a final exam.

Learning Objective/outcome

Upon completing, you will 1) understand the basic concepts and algorithms of machine learning, 2) develop skills of using machine learning libraries for related applications, and 3) acquire
prerequisites for reading research papers and technical reports and implementing cutting-edge machine learning approaches.

Contact Info

Instructor: Kai-Wei Chang

Office Hours: Thu 3pm.

Office Hours Location: https://ucla.zoom.us/j/99875672026?pwd=UXNoZW1QMmJlMERqdCtibThMS2IDQT09

Email: kw+cm146@kwchang.net

Textbooks

While there is not one textbook that covers all the material from this course, readings will come from the following texts:

- *A course in machine learning:* by Hal Daume III, which will be referred to as **CIML** (freely available online) is the primary reference.

For more advanced students, the following are useful:

    - Review notes from Stanford's machine learning course
- Linear algebra
    - Review notes from Stanford's machine learning course
- Optimization
    - Review notes from Stanford's machine learning course
    - Review notes from Stanford's machine learning course

Course Grading Policy (Tentative)

- **Mini quiz on math background (0%)**: This is an online mini quiz that will help you evaluate your background. You can take the exam multiple times.
- **Problem sets (30%)**: There will be periodic problem sets (a.k.a. homework). Questions on the problem sets will include math exercises, programming exercises, and data analyses.
    - We will use gradescope to manage the submission of problem sets.
    - You will also need to submit the source code to bruinlearn
    - Problem sets are due at 11:59 pm on the due date.
    - **Late submissions will not be accepted**
    - All solutions must be clearly written (or typed); unreadable answers will not be graded. We encourage using LaTeX to type out answers.
    - Solutions will be graded on both correctness and clarity. If you cannot solve a problem completely, you will get more partial credit by identifying the gaps in your argument than by attempting to cover them up.
- You can discuss the problems with others. However, you must write up your own solutions. You must also acknowledge all collaborators. Plagiarism is strictly prohibited.
- **Quizzes (20%):** There will be online mini-quizzes every week (except for the first 2 weeks and the last week). One lowest score will be dropped.
- **Exams (Midterm:20% Final: 30 %)**: There is a final exam scheduled in the final week. The exam will be online. You will have a 24-hour window to finish a 3-hour exam. No alternate or make-up exams will be administered, except for disability/medical reasons documented and communicated to the instructor prior to the exam date.

Software

science. For a tutorial, see here

- pytorch: deep learning library.

Forums

Piazza

We will use Piazza for class discussions. Please go to this Piazza website to join the course forum (note: you must use a ucla.edu email account to join the forum). We strongly encourage students to post on this forum rather than emailing the course staff directly (this will be more efficient for both students and staff). Students can use Piazza to:

- Ask clarifying questions about the course material.
- Share useful resources with classmates (so long as they do not contain solutions).
- Look for project partners or other students to form study groups.
- Answer questions posted by other students to solidify your own understanding of the material.

The course Academic Integrity Policy must be followed on the message boards at all times. Do not post or request solutions to problem sets! Also, please be polite.

Gradescope

We will use gradescope to manage and grade problem sets and exams.

- Please follow the link at BruinLearn

Policies

Academic Integrity Policy

Group studying and collaborating on problem sets are encouraged, as working together is a great way to understand new material. Students are free to discuss the homework problems with anyone under the following conditions:

- Students must write their own solutions and understand the solutions that they wrote down.
- Students must list the names of their collaborators (i.e., anyone with whom the assignment was discussed).
- Students may not use old solution sets from this class or any other class under any circumstances, unless the instructor grants special permission.

Although not a formal component of the course grade, attendance is essential for success in this course. Due to the restriction of the classroom, we may not have a full recording of the lectures. We emphatically welcome questions and your active participation in this course will enhance your learning experience and that of the other students.

Regrade requests

Regrade requests for homework and exams must be made through gradescope within one week after the graded homeworks have been released. We reserve the right to regrade all problems for a given regrade request.

Acknowledgments

The course website is based on material developed by Dant Roth, Sriram Sankararaman, Ameet Talwalkar, Andrew Ng, and Fei Sha. Some of the administrative content on the course website is adapted from material from Jenn Wortman Vaughan, Rich Korf, and Alexander Sherstov.

Tentative Schedule (subject to change)

Week	Date	Topic	Quiz	HW	Reading	TA Se (Fri)
Week 0	9/22	Introduction			CIML 2 https://scikit-learn.org/stable/modules/cross_validation.html	Introd Math Backg (Calcu Linear Algert
Week 1	9/27	Nearest Neighbors & Experiment Protocol			CIML 3 https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier	
	9/29	Decision trees			CIML 1 https://scikit-learn.org/stable/modules/tree.html	Explai & Skle
Week 2	10/4	Linear classification/Perceptron			CIML 4, 5 https://scikit-learn.org/stable/modules/linear_model.html	

CM146 Fall 22: Introduction to Machine Learning

Updated automatically every 5 minutes

	10/6	logistic regresion	Quiz 1	HW 1 release	CIML 6.1, 7	(conca conce numer metho
Week 3	10/11	Linear regression / Optimization			CIML 6.3, 9.1 9.2, 9.6, 9.7 https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression	
	10/13	Deep Neural Network	Quiz 3		CIML 10 https://scikit-learn.org/stable/modules/neural_networks_supervised.html https://pytorch.org/tutorials/	Explai / Data Prepa and M practi
Week 4	10/18	Computational Learning Theory			CIML 12	
	10/20	Computational Learning Theory	Quiz 4	HW 2 release	CIML 12	Deep Learn Librar (pyTo
Week 5	10/25	Multiclass classification		Hw1 due	CIML 6.2 https://scikit-learn.org/stable/modules/multiclass.html	
	10/27	Midterm				
Week 6	11/1	Kernel			CIML 11	Explai / VC-Dimer
	11/3	SVM	Quiz 5		https://scikit-learn.org/stable/modules/svm.html	
Week 7	11/8	Ensemble methods			CIML 13 https://scikit-learn.org/stable/modules/ensemble.html	ML applic Case
	11/10	Debug ML System	Quiz 6	HW3 release		
Week 8	11/15	Bayesian Learning		Hw2 due	CIML 9.3-9.5	Explai / Proba Distrib (Norm
	11/17	Naive Bayes	Quiz 7			
Week 9	11/22	Clustering			CIML 15 https://scikit-learn.org/stable/modules/clustering.html	Thank
	11/24	Thanksgiving break				
Week 10	11/29	Expectation-Maximization (EM)			CIML 16	
	12/1	Summary & Practical Concerns		Hw3 due		Final I (Week Week