CSC 480 / 580 – Principles of Machine Learning

Tuesday & Thursday at 5:00pm – 6:15pm
Gould-Simpson, Rm 906

Course Description

Students will learn why machine learning is a fundamentally different way of writing computer programs, and why this approach is often a uniquely attractive way of solving practical problems. Machine learning is all about automatic ways for computers to find patterns in datasets; students will learn both advantages and unique risks that this approach offers. They will learn the fundamental computational methods, algorithms, and perspectives which underlie current machine learning methods, and how to derive and implement many of them. Students will learn the fundamentals of unsupervised and supervised machine learning methods, the computational and quality tradeoffs between different methods, and how to adapt existing methods to fit their own research needs.

Instructor and Contact Information

Instructor:

Jason Pacheco, GS 724
Email: pachecoj@cs.arizona.edu
Webpage: http://pachecoj.com/
Office Hours: Friday, 3-5pm (Zoom link will be provided on Piazza)

Teaching Assistant:

Yinan Li, Email: yinanli@arizona.edu
Office Hours: Monday, 11am-1pm (Zoom link will be provided on Piazza)

Web Information:

Course Webpage: http://pachecoj.com/courses/csc480-580_spring25
D2L CSC 580: https://d2l.arizona.edu/d2l/home/1571901
D2L CSC 480: https://d2l.arizona.edu/d2l/home/1571904
Piazza: https://piazza.com/arizona/spring2025/csc480580
Gradescope CSC 480: Add course via entry code G3Y8E2
Gradescope CSC 580: Add course via entry code WW84YW

Course Format and Teaching Methods

The course will consist of regular in-person lectures. In-class discussion as well as Q&A is encouraged.

Course Objectives

A successful student will be able to implement and explain the limitations of many of the central methods and techniques in machine learning:

- Basic binary classifiers: decision trees, logistic regression
- Supervised vs. unsupervised learning - what's possible in the absence of labels

- Reductions - how to handle imbalanced data; how to build multiclass classifiers
- Practical issues - how to detect overfitting and underfitting; how and when to use feature engineering
- Efficiency issues - how to create classifiers that work well in the presence of large training sets, and large feature sets
- Modern techniques - students will be introduced, via classroom materials and projects, to recent methods in machine learning (this could include, for example, deep learning, reinforcement learning, A/B testing, and multi-armed bandits)

For a more granular description of the learning objectives, see the week-by-week schedule and the description of the assignments below.

Machine Learning is a big field, and there is no way we can cover all of it in one course. With that said, this course covers a large amount of material, and the assignments are a central part of the course. **Students are expected to dedicate a significant amount of time on the course outside of the classroom, especially if they have background deficiencies to make up.**

Expected Learning Outcomes

Expected learning outcomes of the course are:

- To be able to explain the meaning of generalization in machine learning, and why common machine learning algorithms are expected to generalize.
- To be able to list the common supervised and unsupervised learning algorithms and the practical values of each.
- To be able to implement representative machine learning algorithms such as logistic regression, k-nearest neighbors, k-means clustering, etc.
- To be able to identify real-world problems that can be formulated into machine learning problems, perform feature engineering, choose appropriate methods, detect overfitting/underfitting, and evaluate them with statistical significance.
- To be able to explain key challenges in reinforcement learning that supervised/unsupervised learning paradigms do not share.
- To be able to prove the fact that no classifiers can have a smaller test error than the Bayes classifier.
- To be able to mathematically derive the forward-backward algorithm for the hidden Markov model and prove its correctness.

Those taking **CSC 580** have the additional expected outcome:

- To be able to survey a subfield of machine learning, identify gaps in the literature, and develop initial solutions for filling them in.

Makeup Policy for Students Who Register Late

If you register late for this class, contact me as soon as you do. You will be expected to submit all missed assignments within a week of your registration. It is your responsibility to catch up to the class content.

Course Communications

Due to the fact that you have two different D2Ls depending on 480/580, we will use Piazza for the main communications. The main discussion will happen in our piazza forum (link can be found above). However, D2L will be used for making grading available.

Required Texts or Readings

The required textbook is Hal Daumé's Course in Machine Learning (http://ciml.info/), fully and freely available online.

Assignments and Examinations: Schedule/Due Dates

HW0 is due in 7 days and HW1-4 are each due in 10 days.

- 01/16: HW0 (calibration) assigned
- 01/28: HW1 assigned
- 02/13: HW2 assigned
- 03/06: Midterm exam (at the class meeting time)
- 03/18: Project proposal due
- 03/25: HW3 assigned
- 04/08: HW4 assigned
- 05/06: Final project due
- 05/14: Final exam at 6:00pm – 8:00pm

Final Examination

The final exam will happen on May 14th at 6:00pm – 8:00pm. Please see the Final Exam Regulations and Schedules (https://registrar.arizona.edu/faculty-staff-resources/room-class-scheduling/schedule-classes/final-exams)

Grading Scale and Policies

As mentioned above, you will be assessed based on your performance on programming assignments, one final exam, and one project.

The instructing staff will assign grades on a scale from 0 to 100, with the following weights:

- Assignments: 40% (10% × 4)
- Project Proposal: 5%
- Project: 15%
- Midterm Exam: 15%
- Final Exam: 15%
- Participation: 10%

Your final grade in the course will be a direct calculation of the above components and letter grades are as follows:

- 90% or better: A;
- 80% or better: B;
- 70% or better: C;
- 60% or better: D;
- below 60%: E.

For due dates, see "Assignments and Examinations." The homework will be returned to students before the next homework is due. Grading delays beyond promised return-by dates will be announced as soon as possible with an explanation for the delay. As a rule, homework will not be accepted late except in case of documented emergency or illness.

HW0 will not be part of the homework evaluation but will be part of the participation score as it serves as information on the students' background (the participation score will be deducted if the student's submission does not show nontrivial effort to solving it).

By your last day to withdraw, you will know more than 40% of your grade by weight. Before the last day to withdraw (the end of the 10th week), the student will be aware of two homework scores (20%), midterm exam score (15%), and the project proposal assignment (5%). These sum to 40%.

For Those Taking CSC 480

Each homework assignment and exam will have additional questions marked "advanced" that are not required for 400-level credit. In addition, the project requirements will be different between 400- and 500-level credit. Those taking for undergraduate level will not be required to pursue novel research projects. It will be sufficient for undergraduate students to reimplement an existing algorithm, or apply an algorithm to a new dataset and report results. The difference in requirements will lead to different learning experiences as graduate-level students will be required to pursue novel research. As such, graduate-level students will survey the current state of research, identify an area for novelty, and make contributions. Whereas undergraduate-level students will use the project to further their understanding of a topic of their choice.

Incomplete (I) or Withdrawal (W):

Requests for incomplete (I) or withdrawal (W) must be made in accordance with University policies, which are available at https://catalog.arizona.edu/policy/courses-credit/grading/grading-system.

Dispute of Grade Policy:
Dispute of any grading unit (e.g., homework, exam) must be made in a week. Otherwise, the dispute will not be accepted. For the final exam and the final project, the dispute must be made in two days.

Scheduled Topic and Activities

01/16/25	Week 1	Introduction + Course Overview
01/21/25	Week 2	Basics - Decision Trees and Learning Algorithms
01/23/25		Limits - Optimal Bayes Rate Classifier / Overfitting / Underfitting
01/28/25	Week 3	Geometry, NN classifiers, K-Means
01/30/25		Practical Issues - Performance measures, under / over-fitting, CV, pred. Conf
02/04/25	Week 4	Practical Issues (Cont')
02/06/25		Linear Models - Regression
02/11/25	Week 5	Linear Models - Classification
02/13/25		Perceptron
02/18/25	Week 6	Perceptron (Cont'd)
02/20/25		Nonlinear Models

02/25/25	Week 7	Nonlinear Models (Cont'd)
02/27/25		Makeup
03/04/25	Week 8	Makeup
03/06/25		Midterm Exam
03/11/25	Week 9	**Spring Recess - No class**
03/13/25		**Spring Recess - No class**
03/18/25	Week 10	Probability, Naive Bayes, PGMs
03/20/25		Probability, Naive Bayes, PGMs (Cont'd)
03/25/25	Week 11	Neural Networks - Backpropagation
03/27/25		Neural Networks - Convolutional
04/01/25	Week 12	Neural Networks - Autoencoder
04/03/25		Unsupervised Learning
04/08/25	Week 13	Unsupervised Learning (Cont'd)
04/10/25		Ensemble Methods
04/15/25	Week 14	LLMs
04/17/25		LLMs (Cont'd)
04/22/25	Week 15	Reinforcement Learning
04/24/25		Reinforcement Learning (Cont'd)
04/29/25	Week 16	Makeup
05/01/25		Makeup
05/06/25	Week 17	Final Lecture
5/14/2025		Final Exam: 6:00pm - 8:00pm

Classroom Behavior Policy

To foster a positive learning environment, students and instructors have a shared responsibility. We want a safe, welcoming, and inclusive environment where all of us feel comfortable with each other and where we can challenge ourselves to succeed. To that end, our focus is on the tasks at hand and not on extraneous activities (e.g., texting, chatting, reading a newspaper, making phone calls, web surfing, etc.).

Students are asked to refrain from disruptive conversations with people sitting around them during lecture. Students observed engaging in disruptive activity will be asked to cease this behavior. Those who continue to disrupt the class will be asked to leave lecture or discussion and may be reported to the Dean of Students.

Safety on Campus and in the Classroom

For a list of emergency procedures for all types of incidents, please visit the website of the Critical Incident Response Team (CIRT): https://cirt.arizona.edu/case-emergency/overview

Also watch the video available at

https://arizona.sabacloud.com/Saba/Web_spf/NA7P1PRD161/app/me/ledetail;spf-url=common%2Flearningeventdetail%2Fcrtfy000000000003841

University-wide Policies link

Links to the following UA policies are provided here: https://catalog.arizona.edu/syllabus-policies

- Absence and Class Participation Policies
- Threatening Behavior Policy
- Accessibility and Accommodations Policy
- Code of Academic Integrity
- Nondiscrimination and Anti-Harassment Policy

Department-wide Syllabus Policies and Resources link

Links to the following departmental syllabus policies and resources are provided here, https://www.cs.arizona.edu/cs-course-syllabus-policies :

- Department Code of Conduct
- Class Recordings
- Illnesses and Emergencies
- Obtaining Help
- Preferred Names and Pronouns
- Confidentiality of Student Records
- Additional Resources
- Land Acknowledgement Statement

Subject to Change Statement

Information contained in the course syllabus, other than the grade and absence policy, may be subject to change with advance notice, as deemed appropriate by the instructor.