CSC 480 / 580 – Principles of Machine Learning

Gould-Simpson 906, TuTh 9:30-10:45am

Course Description

Machine learning is about automatic ways for computers to collect and/or adapt to data to make predictions, make decisions, or gain insight. It can also be seen as a fundamentally different way of writing computer programs from traditional programming, which is often an attractive way of solving practical problems. Students will learn the fundamental frameworks, computational methods, and algorithms that underlie current machine learning practice, and how to derive and implement many of them. They will also learn both advantages and unique risks that this approach offers.

Course Prerequisites

Linear algebra or equivalent

- *You will need to understand the relationship between linear operators, linear transformations, change of bases, and matrices*
- *We will make repeated use of matrix decompositions such as SVD and eigen-decomposition and their properties.*

Probability and statistics

- *You will need to understand (conditional) expectation and (conditional) independence of random variables and be able to use their key properties such as Bayes' rule and law of total probability.*
- *You will need to understand standard probability distributions including Bernoulli and Gaussian.*

Multivariate calculus or equivalent

- *Recommended: Math 125 (Calc I)*
- *You will need to understand the relationship between the total derivative, the gradient, and how to take advantage of the fact that the differentiation is a linear operator.*

Programming

- *You will need to have a programming maturity at the level of CSC 120. The programming language will be Python.*

Data Science

- *CSC 380 (Principles of Data Science) or equivalent*
- *You will need basic familiarity with the data science workflow and tools*

Instructor and Contact Information

[Chicheng Zhang]

chichengz@cs.arizona.edu

Gould-Simpson 720

Office Hour: TBD

D2L sites:

(480) https://d2l.arizona.edu/d2l/home/1411143

(580) https://d2l.arizona.edu/d2l/home/1411146

Piazza: https://piazza.com/arizona/spring2024/csc480580 access code: 285esihd06d

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

We will use Piazza for the main communications, including important announcements and discussions.

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.

- *1/11: HW0 (calibration) assigned*

- *1/23: HW1 assigned*
- *2/22: HW2 assigned*
- *3/12: Midterm exam (at the class meeting time)*
- *3/19: Project proposal due*
- *3/26: HW3 assigned*
- *4/16: HW4 assigned*
- *5/10: Final project due*
- *5/7: Final exam at 8:00 – 10:00am*

Final Examination

The final exam will happen on May 7th at 8:00-10:00am.

Please see the Final Exam Regulations (https://www.registrar.arizona.edu/courses/final-examination-regulations-and- information) and the Final Exam Schedule, (http://www.registrar.arizona.edu/schedules/finals.htm).

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: 35%
- Project Proposal: 5%
- Project: 20%
- 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 http://catalog.arizona.edu/policy/grades-and-grading-system#incomplete and http://catalog.arizona.edu/policy/grades-and-grading-system#Withdrawal respectively.

Dispute of Grade Policy:

Any grading disputes should be communicated to the professor within one week of having received the grade. The professor will announce in class and on Piazza when grading is complete for each assignment. If a student has not received a grade on a submitted assignment it must be communicated to the adviser within one week of this announcement. If no assignment was submitted it will receive a score of zero.

Honors Credit

Students wishing to contract this course for Honors Credit should e-mail me to set up an appointment to discuss the terms of the contact and to sign the Honors Course Contract Request Form. The form is available at http://www.honors.arizona.edu/honors-contracts

Scheduled Topic and Activities

- *Week 1*
  - *Lecture: Introduction, motivation, course logistics*
  - *Lecture: Basics - Decision Trees, algorithms for learning*
    - *Learning Objectives. Explain the difference between memorization and generalization • Implement a decision tree classifier • Take a concrete task and cast it as a learning problem, with a formal notion of input space, features, output space, generating distribution and loss function.*

- *Week 2*
  - *Lecture: Limits - Optimal Bayes rate and classifier; overfitting and underfitting*
    - *Learning Objectives. Define "inductive bias" and recognize the role of inductive bias in learning • Illustrate how regularization trades off between underfitting and overfitting • Evaluate whether a use of test data is "cheating" or not.*
  - *Lecture: Geometry, nearest-neighbor classifiers, k-means (unsupervised learning preview)*
    - *Learning Objectives • Describe a data set as points in a high dimensional space • Explain the curse of dimensionality • Compute distances between points in high dimensional space • Implement a K-nearest neighbor model of learning • Implement the K-means algorithm for clustering.*

- *Week 3*
  - *Lecture: The perceptron (1/2)*
    - *Learning Objectives. Describe the biological motivation behind the perceptron • Classify learning algorithms based on whether they are error-driven or not • Implement the perceptron algorithm for binary classification • Draw perceptron weight vectors and the corresponding decision boundaries in two dimensions • Contrast the decision boundaries of decision trees, nearest neighbor algorithms and perceptrons • Compute the margin of a given weight vector on a given data set.*
  - *Lecture: The perceptron (2/2)*

- *Week 4*
  - *Lecture: Practical Issues (1/2) - performance measures, underfitting, overfitting, cross validation, prediction confidence via statistical tests and bootstrapping, debugging ML models*
    - *Learning Objectives. Translate between a problem description and a concrete learning problem • Perform basic feature engineering on image and text data • Explain how to use cross-validation to tune hyperparameters and estimate future performance • Compare and contrast the differences between several evaluation metrics.*
  - *Lecture: Practical Issues - (2/2)*

- *Week 5*
  - *Lecture: Bias-variance decomposition, and friends*
    - *Learning Objectives. Understand how classification errors naturally split in approximation error and estimation errors • Understand how error decompositions are useful for debugging.*
  - *Lecture: Reductions (1/3)*
    - *Learning Objectives. Represent complex prediction problems in a formal learning setting • Be able to artificially "balance" imbalanced data • Understand the positive and*

negative aspects of several reductions from multiclass classification to binary classification • Recognize the difference between regression and ordinal regression.

- *Week 6*
  - *Lecture: Reductions (2/3)*
  - *Lecture: Reductions (3/3)*

- *Week 7*
  - *Lecture: Linear Models (1/2)*
    - *Learning Objectives. Define and plot four surrogate loss functions: squared loss, logistic loss, exponential loss and hinge loss • Compare and contrast the optimization of 0/1 loss and surrogate loss functions • Solve the optimization problem for squared loss with a quadratic regularizer in closed form • Implement and debug gradient descent and subgradient descent*
  - *Lecture: Linear Models (2/2)*

- *Week 8*
  - *Lecture: Kernel Methods*
    - *Learning Objectives. Explain how kernels generalize both feature combinations and basis functions • Contrast dot products with kernel products • Implement kernelized perceptron • Derive a kernelized version of regularized least squares regression • Implement a kernelized version of the perceptron • Derive the dual formulation of the support vector machine.*
  - *Midterm exam*

- *Week 9*
  - *Lecture: Probability, Naive Bayes, and Graphical Model Basics (1/3)*
    - *Learning Objectives. Define the generative story for a naive Bayes classifier • Derive logistic loss with an l2 regularizer from a probabilistic perspective.*
  - *Lecture: Probability, Naive Bayes, and Graphical Model Basics (2/3)*

- *Week 10*
  - *Lecture: Probability, Naive Bayes, and Graphical Model Basics (3/3)*
  - *Lecture: Bias and Fairness*
    - *Learning Objectives. Identify how disparity along training/test data can generate bias/unfairness • Understand how a bad choice of metric to optimize can cause bias/unfairness • Identify how careless data collection practices can perpetuate bad decisions • Identify how different assumptions about the world change the way data should be processed for an ML method • Understand how feedback loops can cause arbitrarily bad predictions*

- *Week 11*
  - *Lecture: Neural Networks and Back-Propagation (1/2)*
    - *Learning Objectives. Explain the biological inspiration for multi-layer neural networks • Construct a two-layer network that can solve the XOR problem • Implement the back-propagation algorithm for training multi-layer networks • Explain the trade-off between depth and breadth in network structure • Contrast neural networks with radial basis functions with k-nearest neighbor learning.*

- *Lecture: Neural Networks and Back-Propagation (2/2)*

- *Week 12*
  - *Lecture: Ensembling*
    - *Learning Objectives. Implement bagging and explain how it reduces variance in a predictor • Explain the difference between a weak learner and a strong learner • Derive the AdaBoost algorithm • Understand the relationship between boosting decision stumps and linear classification.*
  - *Lecture: Efficiency*
    - *Learning Objectives. Understand and be able to implement stochastic gradient descent algorithms • Compare and contrast small versus large batch sizes in stochastic optimization • Derive subgradients for sparse regularizers • Implement feature hashing.*

- *Week 13*
  - *Lecture: Unsupervised Learning (1/2)*
    - *Learning Objectives. Explain the difference between linear and non-linear dimensionality reduction • Relate the view of PCA as maximizing variance with the view of it as minimizing reconstruction error • Implement latent semantic analysis for text data • Motivate manifold learning from the perspective of reconstruction error • Understand K-means clustering as distance minimization • Explain the importance of initialization in k-means and furthest-first heuristic • Implement agglomerative clustering • Argue whether spectral clustering is a clustering algorithm or a dimensionality reduction algorithm.*
  - *Lecture: Unsupervised Learning (2/2)*

- *Week 14*
  - *Lecture: Learning Theory (1/2)*
    - *Learning Objectives. Explain why inductive bias is necessary • Define the PAC model and explain why both the "P" and "A" are necessary • Explain the relationship between complexity measures and regularizers • Identify the role of complexity in generalization • Formalize the relationship between margins and complexity*
  - *Lecture: Learning Theory (2/2)*

- *Week 15*
  - *(Reserved for catch-up)*
  - *(Reserved for catch-up)*

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.

Notification of Objectionable Materials

This course will contain material of a mature nature, which may include explicit language, depictions of nudity, sexual situations, and/or violence. The instructor will provide advance notice when such materials will be used. Students are not automatically excused from interacting with such materials, but they are encouraged to speak with the instructor to voice concerns and to provide feedback.

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/common/learningeventdetail/crtfy000000000003560

University-wide Policies link

Links to the following UA policies are provided here, http://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.