

CS 760: Machine Learning Course Overview

Ilias Diakonikolas

University of Wisconsin-Madison

Sept. 8, 2022

Logistics: Lecture Location

•In-person in GRAINGER 2080



Logistics: Enrollment

- •Currently close to capacity
 - Some folks on waitlist may not make it in
 - Sorry $\ensuremath{\mathfrak{S}}$... will be offered again



Logistics: Teaching Team

Instructor: Ilias Diakonikolas

- Location: CS 4386
- Office Hours: TBA / by appointment

TA: Nikolaos (Nikos) Zarifis

• Nikos' OH TBA

Logistics: Content

Three locations:

• 1. Course website: Link from my personal website

http://www.iliasdiakonikolas.org/teaching/Fall22/index.html

•2. Piazza.

https://piazza.com/wisc/fall2022/fa22compsci760001/

- Preferred for questions!
- •3. Canvas



Logistics: Lecture Format

Typically, 75 minutes

- 1-2 breaks with quizzes (ungraded; for understanding only)
- Can also ask questions

We'll post slides on website **before class**

We'll post quizzes after class



Logistics: Assignments & Grades

Homeworks:

- •8 or so, worth 60% total
- Posted after class; due when class starts on due date

Exams:

- Midterm: 20%,
- Final: 20%

Class Setup: Reading

No required textbook, but you should read from the below

- Should all be available online / digital library access
- Will also post articles, papers to read



Class Setup: Background

More on this at the end of class, but

- •Linear algebra (working with data, linear transformations)
- •Calculus (for optimization, convergence, etc.)
- Probability (dealing with noise, sampling)
- **Programming** (for implementation)

Plenty of resources available

 Just need enough experience/mathematical maturity to pick up missing bits

Class Setup: Goals

Two goals:

- Understanding ML
- •Foundation for future research in ML

If you just want to **use** ML, but do not plan to do research, consider taking:

- CS540
- STAT 451
- ECE/CS/ME 532



Class Setup: Goals II

Mini-goals:

- Intuition for each algorithm/model
- •Big picture/ML ecosystem

Examples:

- What types & how much data?
- How hard to train?
- What generalizes best?
- Where is the field going?





Break & Questions

ML Overview: Definition

What is machine learning?

"A computer program is said to learn from experience **E** with respect to some class of tasks **T** and performance measure **P**, if its performance at tasks in **T** as measured by **P**, improves with experience **E**." *Machine Learning*, Tom Mitchell, 1997



ML Overview: Motivation

Why would we do this?

We're building a self-driving car. Could just write down rules
Painful! A lot of cases...

```
/**
 * controls steer ing of the car
 * @param angle
 * @param trim
 */
void steer(float angle, float trim = 0.0) {
   // seems like 360 right 520 left
   PWMPCA9685Device device = new PWMPCA9685Device()
    device.setPWMFrequency(50) //internet says 50hz for servos is optimal
    Servo servo0 = new PCA9685Servo(device.getChannel(channel:1))
   LOG.info("steer angle non corrected:${angle} trim:${trim}")
    if (trim != 0) {
        trim = configTrim
        servo0.setTrim(trim)
    servo0.setInput((angle).toFloat())
   System.out.println("configTrim in service=${configTrim}")
   Thread.sleep(millis: 1000) // important to give time for servo to move
```

ML Overview: Motivation

Why would we do this?

- •We're building a self-driving car. Could just write down rules
 - Painful! A lot of cases...
 - Learn from examples instead



Waymo

Supervised Learning

- •Learning from examples, as above
- Workflow:
 - Collect a set of examples {data, labels}: training set
 - "Train" a model to match these examples
 - "Test" it on new data

Image classification:



indoor



outdoor

Supervised Learning

- •Example: Image classification
- Recall Task/Performance measure/Experience definition
 - Task: distinguish indoor vs outdoor
 - Performance measure: probability of misclassifying
 - Experience: labeled examples

Modality: images



indoor



outdoor

Supervised Learning

• Example: Spam Filtering

- Task: distinguish **spam** vs **legitimate**
- Performance measure: probability of misclassifying
- Experience: labeled examples of messages/emails

							C A Jame G Delete	
		Senty Derward & Arr	hive Delet		From Nature News Alert <nature_news@ealert.nature.com></nature_news@ealert.nature.com>	-W	0/4/12 8:42 AM	
	From fidelity <find-daily@littlesossuscamp.com>12</find-daily@littlesossuscamp.com>	y kepiy - Forward Mark	Unive G Deleo		To Made Concert		9/4/12 8:42 AM	
	Subject \$25k-life-policy-for-\$1-per-month		9/4/12 2:57 P	r	To Mark Craven		Other Actions	
Modality: text	To Mark Craven®		Other Action		Reply B Reply Ist ▼ → Forward Archive A Junk O Delet			
	From Dr. Sanusi Joseph <joseph@yahoo.com>\$</joseph@yahoo.com>	★ Reply → Forward	Not lunk		From Goran Nenadic <g.nenadic@manchester.ac.uk>\$7</g.nenadic@manchester.ac.uk>			
	Subject AFTER A SERIOUS THOUGHT	9/4/12 5:37 P			Subject [BioNLP] New paper on large-scale extraction and contextual	sation of biomolecular events	6/25/12 4:48 F	
	Reply to sanusijoseph@yahoo.cn/2				To bionlp@lists.ccs.neu.edu		Other Action online	
	T₀ undisclosed-recipients: ☆	Other Action:		(→ Reply) → Forward A Junk O Delete			Nature	
		Not lunk		From "Yale, Steven H MD" <yale.steven@marshfieldclinic.org>*</yale.steven@marshfieldclinic.org>				
	From breaking news <find-daily@illinoiscommittee.com>\$</find-daily@illinoiscommittee.com>	forward Archive O Delete	1 1	Subject FW: WGI Demonstration Project Final Report 8/29/12 0		8/29/12 6:52 A		
	Subject green-coffee-bean-study-results:-they-lost-17lbs-in-22-weeks	9/4/12 7:22 PI		To Mark Crave	1	Other Action		
	To Mark Craven	Other Action:		Mark, abs		abstract	abstract	
	Lunk Mail	Not lunk Some						
		one of the		I will work on the draft for the report. Lam still working on adjudicating cases within the MC system				
				with post-hospitalization DVT and PE. I hope to have this done in the next two weeks.				
	is this email not displaying correctly? View it in your browser.							
			illars, the		Thank you.			
	green-coffee-bean-study-results:-th	eV-		Steve				
	green-conce-bean-study-resultsth	ose to be						
	lost-17lbs-in-22-weeks	sorted to		L	available to avanat further recease		Y	
		,			available to support further research.			
		eat did not			Results: Here we present BioContext, an integrated text mining	system		
		but bf their		which extracts, extends and integrates results from a number of tools				
		h out your	n out your		performing entity recognition, biomolecular event extraction and			
	Dr-Oz is calling this a "Miracle-In-A-Bottle".				contextualisation. Application of our system to 10.9 million MEDLINE abstracts and 234,000 open-access full-text articles from PubMed Central			
	The Fresh Green Bean Coffee Diet is being hailed a medical breakthrough in	a weight loss.		l				
	nie neu eren eren eren eren eren eren eren		1					

Supervised Learning

• Example: Ratings/Recommendations

- Task: predict how much a user will like a film
- Performance measure: distance from user's rating
- Experience: previous ratings

Modality: lots







Our best guess for Mark:

Unsupervised Learning

- Data, but no labels. No input/output.
- •Goal: get "something": structure, hidden information, more

• Workflow:

- Collect a set {data}
- Perform some algorithm on it
- **Clustering**: reveal hidden structure



Unsupervised Learning

• Example: Clustering

- Task: produce distinct clusters for a set of data
- Performance measure: closeness to underlying structure
- Experience: available datapoints

Modality: lots



Unsupervised Learning

• Example: Generative Models

- Task: produce artificial images of faces
- Performance measure: photorealism
- Experience: available images

Modality: images



StvleGAN2 (Kerras et al '20)

Reinforcement Learning

- •Agent interacting with the world; gets rewards for actions
- •Goal: learn to perform some activity

• Workflow:

- Create an environment, reward, agent
- Train: modify policy to maximize rewards
- Deploy in new environment
- Controlling aircraft: learn to fly



Reinforcement Learning

• Example: Controlling aircraft

- Task: keep the aircraft in the air
- Performance measure: reward for following trajectory
- Experience: state/action/reward from previous flights

Modality: video/sensor data



Reinforcement Learning

• Example: Playing video games

- Task: play Atari arcade games
- Performance measure: winning/advancing
- Experience: state/action/reward from previous gameplay episodes

Modality: video/sensor data





Break & Questions

Assignment: Reading

For HW1, article by Jordan and Mitchell on course website

REVIEW

Machine learning: Trends, perspectives, and prospects

M. I. Jordan^{1*} and T. M. Mitchell^{2*}

Machine learning addresses the question of how to build computers that improve automatically through experience. It is one of today's most rapidly growing technical fields, lying at the intersection of computer science and statistics, and at the core of artificial intelligence and data science. Recent progress in machine learning has been driven both by the development of new learning algorithms and theory and by the ongoing explosion in the availability of online data and low-cost computation. The adoption of data-intensive machine-learning methods can be found throughout science, technology and commerce, leading to more evidence-based decision-making across many walks of life, including health care, manufacturing, education, financial modeling, policing, and marketing.



achine learning is a discipline focused on two interrelated questions: How can ance when executing some task, through some type of training experience. For example, in learn-

For HW1, self-diagnostic on background. Topics:

- Linear Algebra
- Calculus
- Probability
- Big-O notation
- Basic programming skills



• If these feel very unfamiliar, talk to us

For HW1, self-diagnostic on background. Examples:

Consider the matrix X and the vectors \mathbf{y} and \mathbf{z} below:

$$X = \begin{pmatrix} 9 & 8 \\ 7 & 6 \end{pmatrix} \qquad \mathbf{y} = \begin{pmatrix} 9 \\ 8 \end{pmatrix} \qquad \mathbf{z} = \begin{pmatrix} 7 \\ 6 \end{pmatrix}$$

1. Is X invertible? If so, give the inverse, and if no, explain why not.

2. If $y = \tan(z)x^{6z} - \ln(\frac{7x+z}{x^4})$, what is the partial derivative of y with respect to x?

For HW1, self-diagnostic on background. Examples:

Match the distribution name to its probability density / mass function. Below, $|\mathbf{x}| = k$.

(f) $f(\boldsymbol{x}; \boldsymbol{\Sigma}, \boldsymbol{\mu}) = \frac{1}{\sqrt{(2\pi)^k \boldsymbol{\Sigma}}} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu})\right)$ (g) $f(\boldsymbol{x}; n, \alpha) = \binom{n}{x} \alpha^x (1 - \alpha)^{n-x} \text{ for } \boldsymbol{x} \in \{0, \dots, n\}; 0$ otherwise (h) $f(\boldsymbol{x}; b, \boldsymbol{\mu}) = \frac{1}{2b} \exp\left(-\frac{|\boldsymbol{x} - \boldsymbol{\mu}|}{b}\right)$

- (a) Laplace
- (b) Multinomial
- (c) Poisson
- (d) Dirichlet
- (e) Gamma

- (i) $f(\boldsymbol{x}; n, \boldsymbol{\alpha}) = \frac{n!}{\prod_{i=1}^{k} x_i!} \prod_{i=1}^{k} \alpha_i^{x_i}$ for $x_i \in \{0, \dots, n\}$ and $\sum_{i=1}^{k} x_i = n; 0$ otherwise
- (j) $f(x; \alpha, \beta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$ for $x \in (0, +\infty)$; 0 otherwise
- (k) $f(\boldsymbol{x}; \boldsymbol{\alpha}) = \frac{\Gamma(\sum_{i=1}^{k} \alpha_i)}{\prod_{i=1}^{k} \Gamma(\alpha_i)} \prod_{i=1}^{k} x_i^{\alpha_i 1}$ for $x_i \in (0, 1)$ and $\sum_{i=1}^{k} x_i = 1; 0$ otherwise
- (1) $f(x; \lambda) = \lambda^x \frac{e^{-\lambda}}{x!}$ for all $x \in Z^+$; 0 otherwise

For HW1, self-diagnostic on background. Examples:

Draw the regions corresponding to vectors $\mathbf{x} \in \mathbb{R}^2$ with the following norms:

- 1. $||\mathbf{x}||_1 \le 1$ (Recall that $||\mathbf{x}||_1 = \sum_i |x_i|$)
- 2. $||\mathbf{x}||_2 \le 1$ (Recall that $||\mathbf{x}||_2 = \sqrt{\sum_i x_i^2}$)
- 3. $||\mathbf{x}||_{\infty} \leq 1$ (Recall that $||\mathbf{x}||_{\infty} = \max_{i} |x_{i}|$)

Resources

Probability

• Lecture notes: <u>http://www.cs.cmu.edu/~aarti/Class/10701/recitation/prob_review.pdf</u>

Linear Algebra:

- Short video lectures by Prof. Zico Kolter: <u>http://www.cs.cmu.edu/~zkolter/course/linalg/outline.html</u>
- Handout associated with above video: <u>http://www.cs.cmu.edu/~zkolter/course/linalg/linalg_notes.pdf</u>
 Book: Gilbert Strang. Linear Algebra and its Applications. HBJ Publishers.

Big-O notation:

- <u>http://www.stat.cmu.edu/~cshalizi/uADA/13/lectures/app-b.pdf</u>
 <u>http://www.cs.cmu.edu/~avrim/451f13/recitation/rec0828.pdf</u>

Post others you like!



Thanks Everyone!

Some of the slides in these lectures have been adapted/borrowed from materials developed by Mark Craven, David Page, Jude Shavlik, Tom Mitchell, Nina Balcan, Elad Hazan, Tom Dietterich, Pedro Domingos, Jerry Zhu, Yingyu Liang, Volodymyr Kuleshov, Fred Sala