CS480/680: Introduction to Machine Learning

Lecture 1: Introduction

Hongyang Zhang

Jan 9, 2024

Instructor Team

- **Instructor**:
  - Hongyang Zhang (hongyang.zhang@uwaterloo.ca) (the first 16 lectures)
  - Yaoliang Yu (yaoliang.yu@uwaterloo.ca) (the last 8 lectures)
- **Instructors' Office hour**: every Tuesday 5:20pm-6:20pm, MC 2054
- **TAs**:
  - Haochen Sun (TA head) (h299sun@uwaterloo.ca, design mid-term exam)
  - Ehsan Ganjidoost (eganjidoost@uwaterloo.ca, design final exam)
  - Yanting Miao (y43miao@uwaterloo.ca, design HW2)
  - Alireza Fathollah Pour (a2fathol@uwaterloo.ca, design HW3)
  - Matina Mahdizadeh Sani (m3mahdiz@uwaterloo.ca, design HW4)
  - Shufan Zhang (s693zhan@uwaterloo.ca, design HW1)

Each TA will have their own office hour. You can expect 2-3 OH meetings per week. Please check the course website for TAs' office hour times and locations.

Course Information

- **Times and locations**:
  - Session 001: TTh 8:30am-9:50am, MC 2038
  - Session 002: TTh 10:00am-11:20am, MC 4045
  - Session 003: TTh 4:00pm-5:20pm, MC 2054
- **CS480/680**: CS480 is for undergraduate students; CS680 is for graduate students
- **Website**: `https://watml.github.io/`
  slides, assignments, policy, etc. Please check frequently!
- **Piazza**: https://piazza.com/uwaterloo.ca/winter2024/cs480680/home
  announcements, questions, discussions, etc. Enroll asap!
- **LEARN**: `https://learn.uwaterloo.ca/d2l/home/982358`
  homework submission, grades, etc.
- Require you to attend in person. No video will be recorded!

Prerequisites

- Basic linear algebra, calculus, probability, algorithm
  - CM339 / CS341 or SE 240; STAT 206 or 231 or 241
- Some relevant books and suggested readings on course website
- Coding
  - We will control the use of GPUs in your HWs as small as possible
  - Free GPU use on Google Colab (https://colab.google/)

`https://www.python.org/`

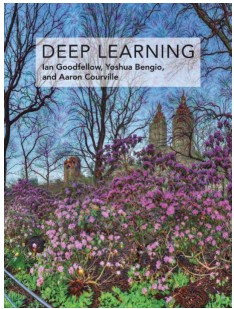
`https://pytorch.org/`

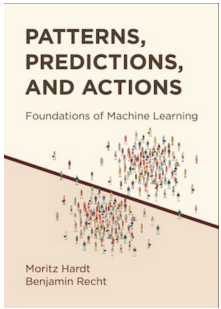
"Coding to programming is like typing to writing. "

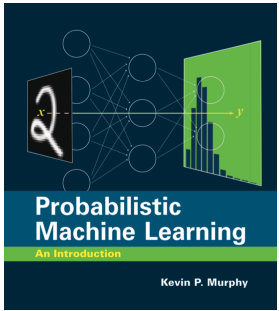
— Leslie Lamport

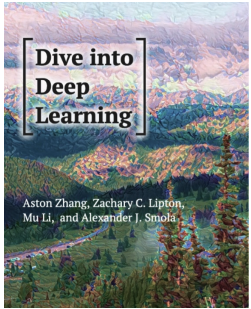
Textbooks

- No required textbook
- Some recommended textbooks:









links available on the course website

Workload

- 24 classes for each session, each lasting 80 mins (including 10-min break in the middle)
- Expect 4 assignments, approx. one every 3 weeks
  - 1/6 (16.66%) weight each
- Mid-term exam: 1/6 (16.66%) weight
- Final exam: 1/6 (16.7%) weight
- Small, constant progress every week
- Submit on LEARN. Submit early and often
  - We do not accept hand-written submission. Typeset using LaTeX is recommended.
  - Submit in the pdf form.
  - We do not accept any submission by email.

Policy

- Do your homeworks **independently and individually**
  - discussion is fine, but no sharing of text or code
  - explicitly acknowledge any source that helps you
- Please be polite and considerate to the TA team
  - TAs are also students like you; they need time to process your request and grade your homeworks
- NO late submissions!
  - Except hospitalization, family urgency; notify beforehand.
  - A formal proof is needed.
  - The proof date should be within 7 days of your homework deadline.
- Late Penalty: Without a proof, your score will be 0 as long as you are late (LEARN submission portal will be closed on time. We DO NOT accept email submission.).

Policy Cont'

- Please inform the TA head Haochen Sun (h299sun@uwaterloo.ca) and provide a screenshot if you have submitted an short-term absence application to Quest for a 2-day ddl extension.
  - Only for CS480 students. You can use this policy once per term (NOT once per course per term).
- Using AI to write homeworks is prohibited. There has been mature tools to detect it (https://gptzero.me/?via=ting&gad_source=1).

Please be considerate!

The three sessions (Sessions 1+2+3) will last for 4.5 hours on every Tuesday and Thursday (plus the office hour afterwards). Sometimes I may lose my voice. So please be considerate.


Questions
?
?
Answers
?

A Brief Introduction of Machine Learning

History of AI

- The 1st AI boom
  - Start: 1950s. Search-based algorithms to solve clearly defined problems
    - For example, Claude Shannon published a detailed analysis of chess playing as search.
  - End: 1970s. Due to disappointments in what AI could deliver

A PROPOSAL FOR THE DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE

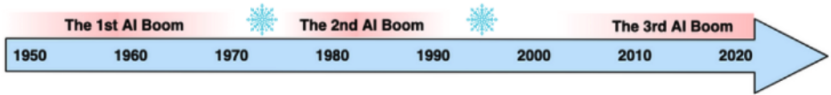
J. McCarthy, Dartmouth College
M. L. Minsky, Harvard University
N. Rochester, I.B.M. Corporation
C.E. Shannon, Bell Telephone Laboratories

August 31, 1955

We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer.

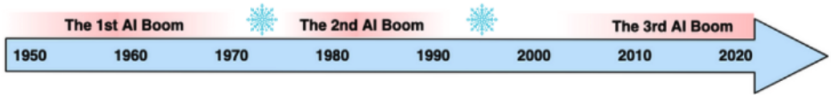
- Key words:
  - Automatic computers
  - Understand a language
  - Self-improvement
  - Abstractions
  - Creativity
  - Achieve all these in only one summer
- Significantly underestimate the difficulty of creating AI

History of AI



- The 2nd AI boom
  - Start: 1980s. Expert systems became popular, about how to represent knowledge.
  - End: 1990s. The AI hype cooled down

History of AI



- The 3rd AI boom
  - Start: 2012. Deep learning freshes new ImageNet competition record
  - Many exciting things: AlphaGo/AlphaStar (game), Alphafold (AI4sci), ChatGPT (AGI), ...

What is Machine Learning (ML)?

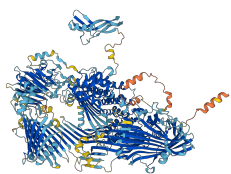
"Machine learning is the field of study that gives computers the ability to learn without being explicitly programmed." — *Arthur Samuel (1959)*

"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." — *Tom Mitchell (1998)*

Machine Learning is Everywhere

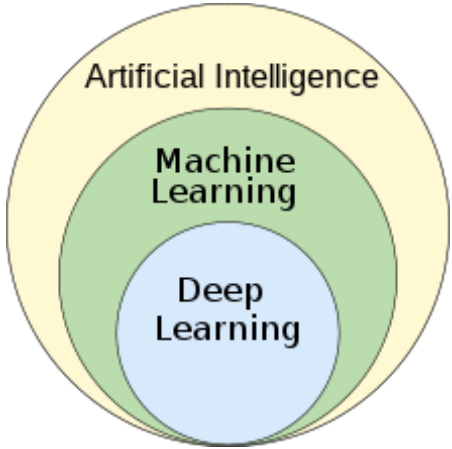
- Machine learning as a service

- Lots of cool applications



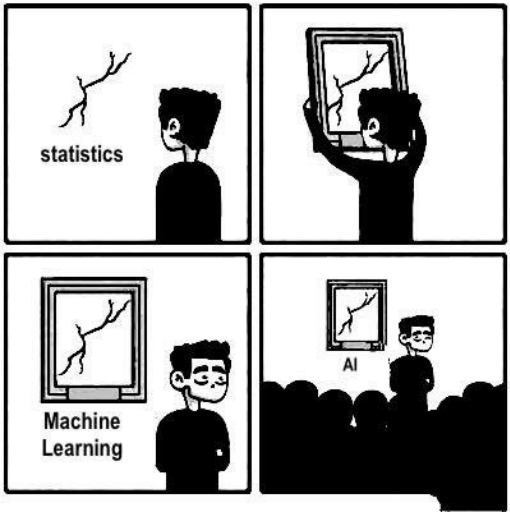
- Excellent for job-hunting

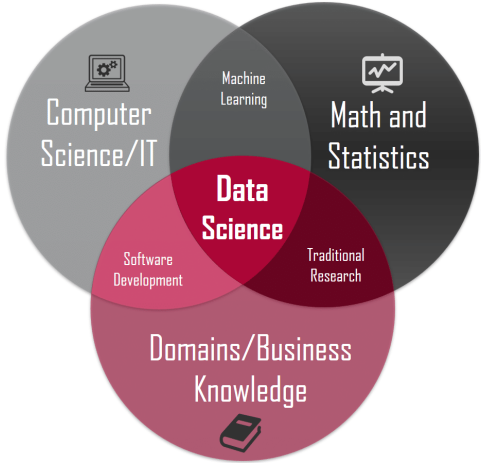
AI, ML and DL



https://en.wikipedia.org/wiki/Machine_learning

Relations with Other Disciplines

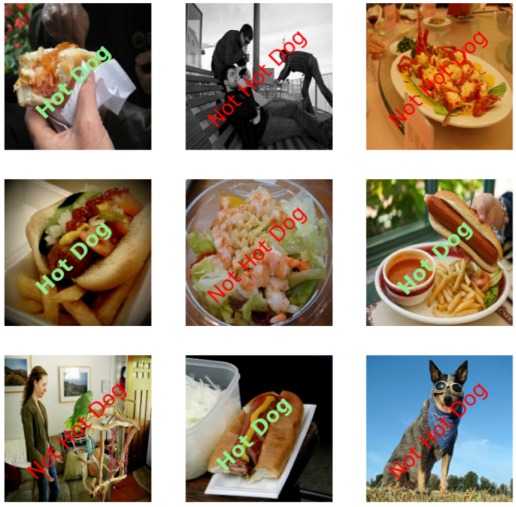




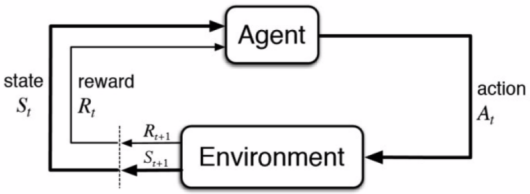
Learning Paradigms

- Supervised learning: teacher/human provides labels (answers)
  - classification: binary, multiclass
  - regression: real-valued, multi-output
  - ranking: recommendation system, Google search
- Unsupervised learning: learning without labels
  - clustering – representation – self-supervised – generative models/AIGC
- Reinforcement learning: machine takes an action; teacher provides rewards
  - control – pricing – games
- Semi-supervised / active learning / etc.

Supervised Learning

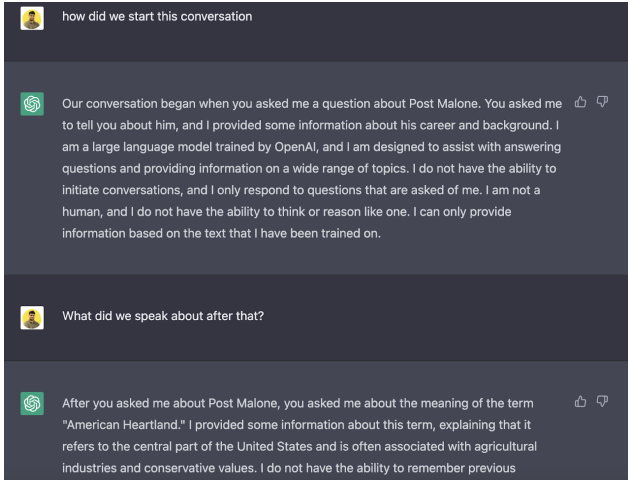


Reinforcement Learning



- Not in this course, but see CS 486/686/885

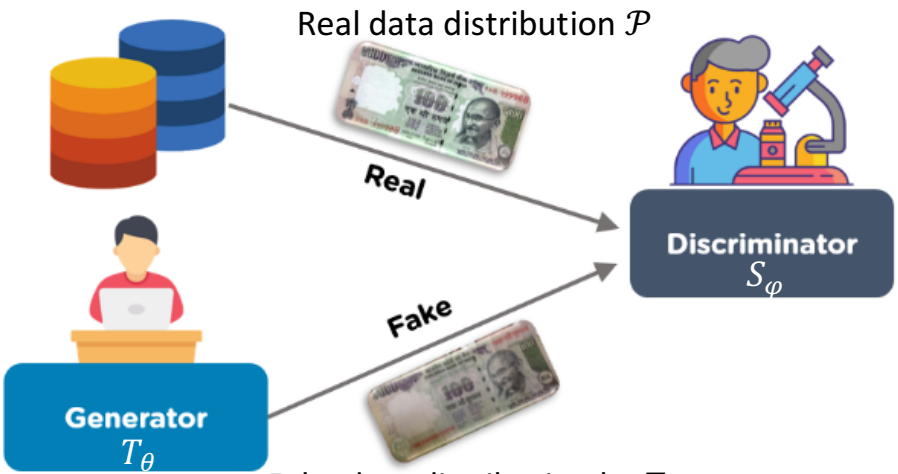
Unsupervised Learning



- Given a prompt, predicting the next word (do not need human annotators to label the texts)

Unsupervised Learning: Generative Adversarial Networks (GAN)

$$\min_{\theta} \max_{\varphi} \mathbb{E}_{\mathbf{x}\sim\mathcal{P}} \log S_\varphi(\mathbf{x}) + \mathbb{E}_{\mathbf{z}\sim\mathcal{N}(0,I)} \log(1 - S_\varphi \circ T_\theta(\mathbf{z}))$$



Unsupervised Learning: Generative Adversarial Networks (GAN)

Unsupervised Learning: Diffusion Models

Focus of ML Research

- **Representation:** how to encode the raw data?
- **Generalization:** how well can we do on unseen data?
- **Interpretation:** how to explain the findings?
- **Complexity:** how much time and space?
- **Efficiency:** how many samples?
- **Privacy:** how to respect data privacy?
- **Robustness:** how to degrade gracefully under (malicious) error and adversarial attacks?
- **Applications**

What You Will LEARN

- Formulate ML problems and recognize pros and cons
- Understand and implement foundational ML models
- Develop and apply ML for new problems on real datasets
- Be ware of potential ethical and safety issues of ML

Outline of the Course

Four modules:

- (I) Classic ML (8 lectures)
- (II) Neural Nets (5 lectures)
- (III) Modern ML Paradigms (4 lectures)
- (IV) Trustworthy ML (5 lectures)

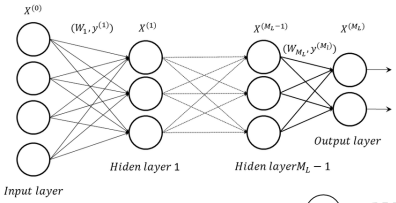
Each homework will correspond to one module.
Warning: Module (I) is all about the foundation of ML and will be more math-heavy.
Modules (II)-(IV) will be application-oriented.

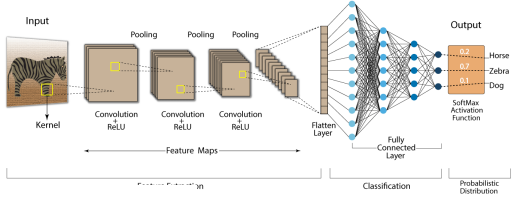
Classic ML

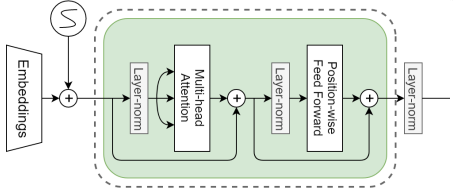
	Date	Category	Topic	Slides	Suggested Readings	Instructor
Lecture 1	Jan 9		**Introduction**	Link	Deep Learning, Section 1	Hongyang Zhang
Lecture 2	Jan 11	Classic ML	**Perceptron**	Link	Patterns, Predictions, and Actions, Page 37	Hongyang Zhang
	Jan 16	Classic ML	**Perceptron - Cont'**	Link	Patterns, Predictions, and Actions, Page 37	Hongyang Zhang
Lecture 3	Jan 18	Classic ML	**Linear Regression**	Link	Probabilistic Machine Learning: An Introduction, Page 363	Hongyang Zhang
Lecture 4	Jan 23	Classic ML	• **Linear Regression - Cont'** • **Logistic Regression**	• Link	Probabilistic Machine Learning: An Introduction, Page 333	Hongyang Zhang
Lecture 5	Jan 25	Classic ML	**Hard-Margin SVM**	Link	The Elements of Statistical Learning, Section 12.3	Hongyang Zhang
Lecture 6	Jan 30	Classic ML	**Soft-Margin SVM**	Link	The Elements of Statistical Learning, Section 12.3	Hongyang Zhang
Lecture 7	Feb 1	Classic ML	• **Soft-Margin SVM - Cont'** • **Reproducing Kernels**	• Link	The Elements of Statistical Learning, Section 12.3	Hongyang Zhang
Lecture 8	Feb 6	Classic ML	**Gradient Descent**	Link	Convex Optimization, Section 9.3	Hongyang Zhang

Neural Nets

Lecture 9	Feb 8	Neural Nets	• **Gradient Descent - Cont'** • **Fully Connected NNs**	• Link	Deep Learning, Section 6	Hongyang Zhang
	Feb 13	Neural Nets	**Fully Connected NNs - Cont'**	Link	Deep Learning, Section 6	Hongyang Zhang
Lecture 10	Feb 15	Neural Nets	**Convolutional NNs**	Link	Deep Learning, Section 9	Hongyang Zhang
	Feb 27	Neural Nets	**Convolutional NNs - Cont'**	Link	Deep Learning, Section 9	Hongyang Zhang
No class	Feb 29	-	**Mid-term Exam**	-	-	Hongyang Zhang
Lecture 11	March 5	Neural Nets	**Transformer**	Link	• "Attention Is All You Need". Vaswani et al. 2017 link	Hongyang Zhang

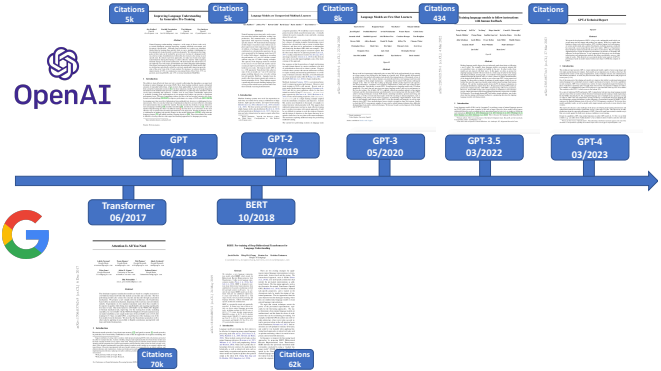






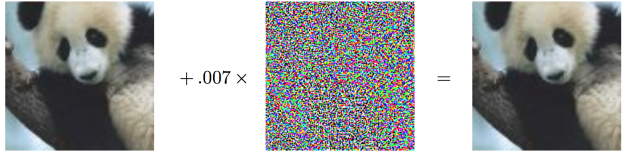
Modern ML Paradigms

Lecture 12	March 7	Modern ML Paradigms	**Large Language Models**	Link	• "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding". Devlin et al. 2018 link • (GPT-1) "Improving Language Understanding by Generative Pre-training". Radford et al. 2018 link • (GPT-2) "Language Models are Unsupervised Multitask Learners". Radford et al. 2019 link • (GPT-3) "Language Models are Few-Shot Learners". Brown et al. 2020 link • (GPT-3.5) "Training Language Models to follow Instructions with Human Feedbacks". Ouyang et al. 2022 link • (GPT-4) "GPT-4 Technical Report". OpenAI 2023 link • (talk by Andrej Karpathy) State of GPT link	Hongyang Zhang
Lecture 12	March 12	Modern ML Paradigms	**GANs**	Link	• "Generative Adversarial Networks". Goodfellow et al. 2014 link	Yaoliang Yu
Lecture 13	March 14	Modern ML Paradigms	**Flow**	Link	TBD	Yaoliang Yu
Lecture 14	March 19	Modern ML Paradigms	**Self-Supervised Learning**	Link	• "A Simple Framework for Contrastive Learning of Visual Representations". Chen et al. 2020 link • "Momentum Contrast for Unsupervised Visual Representation Learning". He et al. 2020 link	Yaoliang Yu



Trustworthy ML

Lecture 15	March 21	Trustworthy ML	**Evasion Attacks**	Link	• (White-box) "Reliable evaluation of adversarial robustness with an ensemble of diverse parameter-free attacks". Croce et al. ICML 2020. link • (White-box) "Obfuscated gradients give a false sense of security: Circumventing defenses to adversarial examples". Athalye et al. ICML 2018. link • (White-box) "DeepFool: a simple and accurate method to fool deep neural networks". Moosavi-Dezfooli et al. CVPR 2016. link • (Black-box) "Square Attack: a query-efficient black-box adversarial attack via random search". Andriushchenko et al. ECCV 2020. link • (Black-box) "Decision-Based Adversarial Attacks: Reliable Attacks Against Black-Box Machine Learning Models". Brendel et al. ICLR 2018. link	Yaoliang Yu
Lecture 16	March 26	Trustworthy ML	**Robustness**	Link	• "Towards Deep Learning Models Resistant to Adversarial Attacks". Madry et al. ICLR 2018 link • "Theoretically Principled Trade-off between Robustness and Accuracy". Zhang et al. ICML 2019 link	Yaoliang Yu
Lecture 17	March 28	Trustworthy ML	**Privacy**	Link	DifferentialPrivacy.org	Yaoliang Yu
Lecture 18	April 2	Trustworthy ML	**Fairness**	Link	TBD	Yaoliang Yu
Lecture 19	April 4	Trustworthy ML	• **Other Threats** • **Course Review**	Link	• (Physical) "Robust Physical-World Attacks on Deep Learning Models". Eykholt et al. CVPR 2018 link • (Physical) "Adversarial examples in the physical world". Kurakin et al. ICLR 2017 link • (Physical) "Synthesizing Robust Adversarial Examples". Athalye et al. ICML 2018 link • (Physical) "Fooling automated surveillance cameras: adversarial patches to attack person detection". Thys et al. CVPR 2019 Workshop link • (Poisoning) "Poison Frogs! Targeted Clean-Label Poisoning Attacks on Neural Networks". Shafahi et al. NeurIPS 2018 link • (Poisoning) "Trojaning Attack on Neural Networks". Liu et al. NDSS 2018 link • (Poisoning) "Hidden Trigger Backdoor Attacks". Saha et al. AAAI 2020 link • (Poisoning) "Deep Partition Aggregation: Provable Defense against General Poisoning Attacks". Levine et al. ICLR 2021 link	Yaoliang Yu




Questions
?
?
Answers
?