Epistemology of Machine Learning (80326/80626)
Course syllabus: Spring 2021

Course time: Tuesdays & Thursdays 12:20--01:40PM

Course location:
https://cmu.zoom.us/j/91713963052?pwd=Rng0Z1VnYk92NEYrWGtWZ3gxaGtnZz09

Instructors:

Kevin Kelly (kk3n@andrew.cmu.edu)
Office: 135K Baker Hall
Phone: 412.268.8567
Zoom office:
https://cmu.zoom.us/j/2374855184
Office Hours: Tuesday 5:30-6:30PM
(other times by appointment)

Kun Zhang (kunz1@cmu.edu)
Office: 161B Baker Hall
Phone: 412.268.8573
Zoom office:
https://cmu.zoom.us/j/8214572323
Office Hours: Wednesday 10:00—11:00AM
(other times by appointment)

Description

Learning is hard. Programming a computer is also hard. Better that computers learn on their own from data how best to serve us. That is the goal of machine learning (ML), which is arguably the most successful branch of artificial intelligence. The very idea raises some natural, fundamental questions. What, exactly, is the goal of learning? Is it maintenance of consistency among our beliefs (Bayesian statistics), or is it a matter of estimating or predicting quantities in nature (frequentist statistics)? Is the goal prediction or truth; control or understanding; actual fact or necessary causal relations? Are predictions expected to be reliable beyond the narrow circumstances of training? Can learning be guaranteed to succeed by a specified time? Does success entail some detectable mark or sign of success that can be used to terminate the learning procedure? What roles do causality and simplicity play in learning, and how?

Those questions arise spontaneously for the reflective ML researcher, but they also cross over into the foundations of statistics, the philosophy of science, and epistemology, the traditional philosophical study of the nature of knowledge and justified belief. This class provides an introduction to the relevant philosophical/foundational issues underlying ML research. It involves both philosophical reflection and concrete familiarity with ML methods. A unifying theme is that strong definitions of successful learning imply correspondingly strong limits on what can be learned. The only prerequisite is one solid course in probability or statistics.

Course objectives

As an outcome of this course, participants are expected to:

- Become acquainted with some basic statistical and machine learning problems and methods.
- Become acquainted with the fundamental, epistemological distinctions: belief vs. knowledge, reliability vs. coherence, induction vs. deduction, realism vs. anti-realism, and explanation vs. prediction.

- Be able to apply the fundamental epistemic distinctions aptly to statistical and machine learning problems and methods.
- Understand the difference between frequentist and Bayesian statistics, and how it relates to the fundamental epistemological distinctions.
- Particular attention will be placed on the essential roles of simplicity and causality in statistics and machine learning.
- Other machine learning courses focus on what to *do*. This course focuses on the deeper question of why you *should*. In particular, we will focus on what it even means to learn well or poorly.

Who can attend

Students are expected to have taken one course in probability theory, introductory statistics, or machine learning.

Course materials

Reading materials will be available online or in Canvas.

Grading

Class participation is essential for success, since there is no textbook that covers exactly the material we will present. You need to prepare for the class discussion by thinking through the ideas in the reading. To get credit for your preparation, you will submit reading questions for each reading assignment. The assignments will be posted on the Canvas class site. If the reading questions are not posted at the beginning of the corresponding class, you will receive a 10% deduction for that reading assignment (Canvas automatically time-stamps all submissions). The reading questions will constitute 1/3 of your total grade.

You will also write a short midterm essay and a short final essay, in which you explain how to apply epistemological concepts to machine learning problems or tasks. Each essay will constitute 1/3 of your total grade. Graduate students will write 8-page papers.

Remarks: We will grade undergraduate students on a curve entirely separately from graduate students. We take account of students' diverse backgrounds when we grade the essays.

Class schedule (subject to change: please refer to the updated version on Canvas)

The course is divided into seven Parts, as follows.

Part I. Introduction to ML: Definition and typical tasks [2 weeks]

- **2/2 (Tuesday) Intro to Class**
  - Personal introductions.
  - Class mechanics and expectations.
  - Class concept (epistemological distinctions applied to ML problems and techniques and their inspirations in ML).

- **2/4 (Thursday) Plato on learning.** (The *Meno*).
  ✓Reading: *The Meno* (http://www.gutenberg.org/files/1643/1643-h/1643-h.htm);

- **2/9 (Tuesday) and 2/11 (Thursday): Intro to AI and ML.**
  - Definition of machine learning and artificial intelligence.
  - Formulation of learning problems: supervised learning, unsupervised learning, and reinforcement learning.
  - How to make ML systems adaptive and intelligent?
  - Introduction to "simplicity" in ML.
  - Causality problems: why causality, identification of causal *effects, counterfactual* reasoning, and causal *discovery.*

    ✓Reading: Sections "Overview," "History and relationships to other fields," and "Approaches" of the Wikipedia page: https://en.wikipedia.org/wiki/Machine_learning.

Part II. Five fundamental distinctions in epistemology [2.5 weeks]

- **2/16 (Tuesday) Belief vs. Knowledge**
  ✓Reading: https://plato.stanford.edu/entries/knowledge-analysis/.

- **2/18 (Thursday) Reliability vs. Coherence**
  ✓Reading: https://plato.stanford.edu/entries/reliabilism/;
  ✓Reading: https://plato.stanford.edu/entries/justep-coherence/.

- **2/23 (Tuesday) *Break Day; No Class***
-
- **2/25 (Thursday) Induction vs. Deduction**
  ✓Reading: https://plato.stanford.edu/entries/induction-problem/#HumeProb.

- **3/2 (Tuesday) Realism vs. Anti-realism**
  ✓Reading: https://plato.stanford.edu/entries/scientific-realism/.

- **3/4 (Thursday) Explanation vs. Prediction**
  ✓Reading: https://plato.stanford.edu/entries/scientific-explanation/.

Part III. The Epistemology of Statistics

- **3/9 (Tuesday) The mathematics of probability I.**
  - Probability axioms
  - Definition of conditional probability
  - Total probability and Bayes' theorem
  - Discrete and continuous variables
  - Probability distributions

- Typical distributions

- **3/11 (Thursday) The mathematics of probability II.**
  - Expectation.
  - Independence and conditional independence.
  - Gaussian distribution.
  - Central limit theorem and Crámer decomposition theorem.

- **3/16 (Tuesday) Frequentist statistics I**
  - Probabilities as frequencies or chances.
  - Testing (five distinctions).
  - Interval estimation (five distinctions).
  - Maximum likelihood estimation and test.

- **3/18 (Thursday) Frequentist statistics II**
  - Point estimation: bias, variance, efficiency (five distinctions).
  - Regression and prediction (five distinctions).

    ✓Reading: 1. https://plato.stanford.edu/entries/statistics/ (section on frequentism)
    2. Chapter on maximum likelihood estimation of "Probability and Statistical Inference" (by R. V. Hogg, E. A. Tanis, and D. L. Zimmerman; available on Canvas)

- **3/23 (Tuesday) Bayesian statistics**
  - Degrees of belief and betting.
  - Coherent degrees of belief and Dutch Book.
  - Bayesian conditioning and learning (five distinctions).
  - Bayesian reconstructions of testing, interval estimation, point estimation.

    ✓Reading: https://plato.stanford.edu/entries/statistics/ (section on Bayesianism).

Part IV. The Epistemology of Supervised Machine Learning

- **3/25 (Thursday) Supervised Learning: prediction as regression.**
  - Parametric regression and confidence bounds.
  - Non-parametric regression.
  - Uniformity of nature: kernel regression, gaussian process regression.

- **3/30 (Thursday) Classification:**
  - Francis Bacon's inductive method and its connection to modern classification/regression
  - Advanced methods for supervised learning
  - Vapnik-Chervonenkis dimension and PAC learning.

    ✓Reading:
    https://www.aub.edu.lb/fas/cvsp/Documents/Bacon%C2%B9s%20inductive%20method,%20example%20of%20heat.pdf.

Part V. Ockham's Razor in ML [1.5 week]

- **4/1 (Thursday) Bayesian razors**
  - Bayesian Ockham's razor and Bayes Information Criterion (BIC).
  - Kolmogorov complexity and Minimum Description Length (MDL).

- **4/6 (Tuesday) Frequentist razors**
  - The "over-fitting argument" and Akaike Information Criterion (AIC), cross validation.
  - Structural Risk Minimization (SRM).

- **4/8 (Thursday) Frequentist razors**
  - Plato's razor from truth preservation.

- **4/13 (Tuesday)** ***No Class***

Part VI. Causality and ML [2.5 week]

- **4/15 (Thursday) Causality: concepts**
  - Causality and interventions
  - Properties of causal systems: modularity
  - Causal graphical models

    ✓**Reading**: Pages 1-6 of "Causal discovery and inference: concepts and recent methodological advances" (By Spirtes & Zhang), Applied Informatics, 2016

- **4/20 (Tuesday) Causality: typical problems**
  - Three typical problems in AI: prediction, identification of causal effects, and counterfactual reasoning.

    ✓**Reading**: Chapter 1 of Pearl's book ("Causality: Models, Reasoning, and Inference," published in 2000; available on Canvas).

- **4/22 (Thursday) Learning causality from data with ML ideas**
  - Why is it possible to learn causality?
  - Constraint-based methods for causal discovery: PC and FCI.

    ✓Reading: Chapters 5.4.1 & 5.4.2 of the SGS book ("Causation, Prediction, and Search"); available on Canvas.

- **4/27 (Tuesday)** Functional causal model-based methods: Linear, non-Gaussian, Acyclic Model (LiNGAM).
  - Gaussian distribution vs. non-Gaussian distribution
  - Asymmetry in variables in linear, non-Gaussian cases
  - Independent component analysis for estimating LiNGAM

    ✓Reading: "A linear non-Gaussian acyclic model for causal discovery" (by Shimizu et al.), Journal of machine learning research, 2006

Part VII. Epistemology of advanced learning problems [2 week]

- **4/29 (Thursday) Unsupervised learning and deep learning**

- Clustering
- A causal view of unsupervised learning and semi-supervised learning

- **5/4 (Tuesday) Deep learning**
  - What it is.
  - What it does.
  - Representation learning and Carnap's "The Logical Structure of the World"

- **5/6 (Thursday)5/4 (Tuesday) Adversarial vulnerability and general-purpose AI**
  - Adversarial vulnerability: what it is and why it happens
  - Possible ways to deal with adversarial vulnerability
  - Transfer learning and image-to-image translation
  - Fairness in ML
  - Towards general-purpose AI
    ✓Reading: "Domain Adaptation As a Problem of Inference on Graphical Models," by Zhang et al., NeurIPS 2020

Additional information

Use of Zoom in the class

In our class, we will be using Zoom for synchronous (same time) sessions. The link is https://cmu.zoom.us/j/91713963052?pwd=Rng0Z1VnYk92NEYrWGtWZ3gxaGtnZz09 . *Please make sure that your Internet connection and equipment are set up to use Zoom and able to share audio and video during class meetings. (See this page from Computing Resources for information on the technology you are likely to need.) Let me know if there is a gap in your technology set-up (email: kk3n@andrew.cmu.edu* or *kunz1@cmu.edu) as soon as possible, and we can see about finding solutions.*

Sharing video: In this course, being able to see one another helps to facilitate a better learning environment and promote more engaging discussions. Therefore, our default will be to expect students to have their cameras on during lectures and discussions. However, I also completely understand there may be reasons students would not want to have their cameras on. If you have any concerns about sharing your video, please email us as soon as (email: kk3n@andrew.cmu.edu or *kunz1@cmu.edu) and we can discuss possible adjustments. Note: You may use a background image in your video if you wish; just check in advance that this works with your device(s) and internet bandwidth.*

Diversity statement

We must treat every individual with respect. *We are diverse in many ways, and this diversity is fundamental to building and maintaining an equitable and inclusive campus community. Diversity can refer to multiple ways that we identify ourselves, including but not limited to race, color, national origin, language, sex, disability, age, sexual orientation, gender identity, religion, creed, ancestry, belief, veteran status, or genetic information. Each of these diverse identities, along with many others not mentioned here, shape the perspectives our students, faculty, and staff bring to our campus. We, at CMU, will work to promote diversity, equity and inclusion not only because* *diversity fuels excellence and innovation, but because we want to pursue justice. We acknowledge our imperfections while we also fully commit to the work, inside and outside of our classrooms, of building and sustaining a campus community that increasingly embraces these core values.*

Each of us is responsible for creating a safer, more inclusive environment.

Unfortunately, incidents of bias or discrimination do occur, whether intentional or unintentional. They contribute to creating an unwelcoming environment for individuals and groups at the university. Therefore, the university encourages anyone who experiences or observes unfair or hostile treatment on the basis of identity to speak out for justice and support, within the moment of the incident or after the incident has passed. Anyone can share these experiences using the following resources:

- ***Center for Student Diversity and Inclusion:*** *csdi@andrew.cmu.edu, (412) 268-2150*
- ***Report-It online anonymous reporting platform:*** *reportit.net username**:** tartans password**:** plaid*

All reports will be documented and deliberated to determine if there should be any following actions. Regardless of incident type, the university will use all shared experiences to transform our campus climate to be more equitable and just.

Course Policies:
Remember: If you registered for this class, you have until November 9th to change your grade in this course from a letter grade to a Pass/Fail grade.

Cheating and Plagiarism:
It is the responsibility of each student to be aware of the university policies on academic integrity, including the policies on cheating and plagiarism. This information is available at http://www.cmu.edu/academic-integrity.

Disability

If you have a disability and have an accommodations letter from the Disability Resources office, I encourage you to discuss your accommodations and needs with me as early in the semester as possible. I will work with you to ensure that accommodations are provided as appropriate. If you suspect that you may have a disability and would benefit from accommodations but are not yet registered with the Office of Disability Resources, I encourage you to contact them at access@andrew.cmu.edu.

Student Well-Being

Take care of yourself. Do your best to maintain a healthy lifestyle this semester by eating well, exercising, avoiding drugs and alcohol, getting enough sleep and taking some time to relax. This will help you achieve your goals and cope with stress.

All of us benefit from support during times of struggle. There are many helpful resources available on campus and an important part of the college experience is learning how to ask for help. Asking for support sooner rather than later is almost always helpful.

If you or anyone you know experiences any academic stress, difficult life events, or feelings like anxiety or depression, we strongly encourage you to seek support. Counseling and Psychological Services (CaPS) is here to help: call 412-268-2922 and visit their website at http://www.cmu.edu/counseling/. Consider reaching out to a friend, faculty or family member you trust for help getting connected to the support that can help.