Introduction to Machine Learning

Michael Yao, Allison Chae

Learning Objectives

1. **Describe** what an algorithm is and how they are used in both clinical medicine and everyday life.

2. **Describe** what it means to learn and how learning applies to machine learning.

3. **Identify** key applications of machine learning and when computational tools can be helpful (and potentially harmful) for patient care.

Food for Thought

1. Are there tasks in healthcare for which automated methods–such as computer algorithms and artificial intelligence (AI)–should never be used? Why or why not?

2. Does understanding ***how*** automated methods arrive at their predictions change any of your answers to question 1?

💡 What if automated methods perform the task on par with humans? What if they perform *better than* humans?

Introduction to Machine Learning

What is an algorithm?

An **algorithm** is any function that computes an output from an input. We already use algorithms in everyday life and in clinical medicine. For example, here is an algorithm that you might use to determine when to walk to JMEC based on 3 different variables:

```
y = some_algorithm_for_when_to_walk_to_jmec(
    how_long_does_it_typically_take_to_walk_to_campus,
    how_much_sleep_did_I_get_last_night,
    is_class_mandatory
)
```

where `y` is when you decide to walk to campus.

Some algorithms can be written down exactly. For example, compute the anion gap given patient values, or compute the MAP of a patient given their blood pressure.

Other algorithms are harder to express on paper. For example, how to run a code or how to determine whether to admit a patient or not.

Computers can run algorithms that can be written down exactly. But how can we teach them how to run algorithms that are hard to express? To answer this question, let's reflect on how we as students learn algorithms that might be hard to express.

Learning by Observing

Computers can learn by observation, much like how medical students learn! Consider some of the following scenarios:

A Database of Genomes

During your clinical research year, your advisor gives you a large dataset of many different patient genomes. By analyzing

this dataset, we try to gain insights into which genes make individuals unique, and which ones all patients share in common.

A Randomized Control Trial

Your research mentor is impressed with your analysis and gives you a new project: investigating if a new drug **abastatin** lowers patient cholesterol levels. He gives you a large dataset of anonymized patient data containing two variables: whether the patient was given abastatin or placebo (x), and whether the patient had a reduction in their cholesterol levels (y). By analyzing this dataset, we try to learn whether or not abastatin is an effective drug for hypercholesterolemia.

A Patient with Sepsis

A 52 year-old male presents with acute-onset altered mental status and fever. Vitals are notable for BP 90/60 and T 103.4. We can denote the patient as a variable x consisting of all of the relevant attributes of the patient: their HPI, past medical history, current lab values and vitals, etc.

On our first day as a medical student, we might not know what to do with this patient. Do we admit them and start them on IV antibiotics? Do we call a neurology consult? Do we just send the patient home? Each of these clinical interventions can be thought of as an action a that we can take to try to help the patient get better.

After observing a patient x and performing an action a, we monitor the patient to see if they improve. The patient's outcome can be denoted as a variable y (for example, $y = 0$ if the patient deteriorates and $y = 1$ if the patient gets better). We observe the clinical outcome y, and use it to learn a better algorithm so that next time we see a similar patient, we can take a better action that leads to a more favorable outcome.

Over the course of medical school, we see hundreds (if not thousands) of tuples (x, a, y) through clerkships, sub-Is, exams, and UWorld, and use this dataset of patient-action-outcome observations to learn *hard-to-write-down algorithms* for choosing the best clinical intervention a given a patient x to maximize the outcome y.

In other words, **we learn by observation**.

❶ For those of you with a machine learning background, **A Database of Genomes** is an *unsupervised learning problem*, **A Randomized Control Trial** is a *supervised learning problem*, and **A Patient with Sepsis** is a *reinforcement learning problem*. You can learn more about each of these types of machine learning problems here!

What does it mean to "learn"?

No patient is exactly identical to any other patient, including the patients that you learn from. If all you can do is regurgitate the dataset you learned from, this is not learning! Put simply...

> Learning = Generalization
>
> After observing the many different patient cases and outcomes, we want to be able to generalize to new patients in the future, such that we know what to do as clinicians for future, previously unknown patients.

Machine Learning as a Framework

Machine learning (ML) uses the *exact same* framework of **learning through observation** to learn hard-to-write-down algorithms from data as ***exact steps*** that a computer can execute.

The fundamental goal of machine learning is to learn **hard-to-write-down** algorithms from past observations to hopefully make accurate predictions for future observations.

> What problems might cause algorithms to generalize poorly to new patients?
>
> There are a number of reasons. Here are a couple:
>
> 1. **New patients are very different from the patients used to learn the algorithm.** For example, society guidelines developed in the United States may lead to substandard care if implemented directly in another country like Korea or Nigeria. This is because the prevalence (and potentially pathophysiology) of diseases may differ between different areas of the world.
> 2. **The algorithm has too many inputs and learns relationships between inputs and out-**

ⓘ Just like how we all have different mnemonics and mental maps on how to approach clinical reasoning, the exact steps in the algorithm that ML learns may not be the same as the steps that clinicians learn! This is an important problem that researchers are still trying to solve.

puts that are not true. If you observe enough features, you may find "correlations" that end up just being due to random chance. For example, if we include "ice cream sales" as an input to an algorithm to predict drowning death rate, we may incorrectly learn that ice cream causes drowning. (This is also referred to as *spurious correlations*)

3. **The algorithm is learned from too small of a patient population.** If we only observe five septic patients and notice that all of them eat candy and miraculously get better, we might conclude that sugar cures sepsis. Generalizing before we've seen enough observations can lead to incorrect conclusions.

What other problems did you think of? Are there any ways to fix the problems that we've identified?

When can machine learning be a helpful tool?

Consider the following example cases. Would you want to use machine learning in each of these cases?

1. Given patient lab values and health record data, ML predicts the age of a patient.

No, it's easy to just look up the age of a patient from the patient chart.

2. Given patient blood pressure values, ML predicts the patient's MAP.

No, computing MAP is an easy-to-write-down algorithm.

3. Given patient lab values, imaging data, genomic data, and other attributes, ML predicts whether the patient is at risk for Huntington's disease (a disease with no known cure).

No, even if ML derives an algorithm for this task, there is nothing actionable that we can do about it.

4. Given patient bowel sound recordings, ML predicts the probability a patient has an SBO.

No, we don't have any datasets of patient bowel sounds mapping to the presence/absence of an SBO, so there are no prior observations for ML to learn from.

5. Given patient lab values, ML predicts the probability a patient has ribose-5-phosphate isomerase (RPI) deficiency, the second rarest disease in the world.

No, we don't have nearly enough observations of patients with RPI deficiency. Even if we did have enough data to learn this algorithm, we also would rarely/never even need to use this algorithm.

In summary, ML is useful for tasks that are

1. hard-to-write-down;
2. associated with a lot of prior observations; and
3. can lead to actionable utility for patients and/or clinicians by automating hard, repetitive, and/or common tasks.

Evidence-Based Medicine Discussion

Should AI be used to improve access to mental health resources?

1. Overview Article

Stade EC, Stirman SW, Ungar LH, Boland CL, Schwartz HA, Yaden DB, Sedoc J, DeRubeis RJ, Willer R, Eichstaedt JC. Large language models could change the future of behavioral healthcare: A proposal for responsible development and evaluation. npj Mental Health Res 3(12). (2024). doi: 10.1038/s44184-024-00056-z. PMID: 38609507

tl;dr: Large language models (LLMs) are AI tools that can read, summarize, and generate text. Early research efforts are investigating the applications of LLMs for psychotherapy. These tools may improve psychotherapy care delivery and access to mental health resources. However, poor outcomes or ethical transgressions from clinical LLMs could also harm patients.

2. Yes, AI is more empathetic than physicians.

Ayers JW, Poliak A, Dredze M, Leas EC, Zhu Z, Kelley J, Faux DJ, Goodman AM, Longhurst CA, Hogarth M, Smith DM. Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. JAMA Intern Med 183(6): 589-96. (2023). doi: 10.1001/jamainternmed.2023.1838. PMID: 37115527

tl;dr: Cross-sectional study using 195 randomly drawn public patient questions from Reddit's r/AskDocs forum. Authors compared physician and chatbot responses to these questions. The chatbot responses were preferred over physician responses and rated significantly higher for both quality and empathy. AI assistants may be able to aid in drafting responses to patient questions.

3. No, AI is too slow to appropriately escalate mental health risk scenarios.

Heston TF. Safety of large language models in addressing depression. Cureus 15(12): e50729. (2023). doi: 10.7759/cureus.50729. PMID: 38111813

tl;dr: Cross-sectional study evaluting 25 conversational AI chatbots specifically designed for mental health counseling. Each chatbot was evaluated twice via highly structured patient simulations designed to assess if (1) the chatbot can escalate suicide risk based on Patient Health Questionnaire (PHQ-9) scores; and (2) the chatbot can recognize suicidality in conversation. Few chatbots included crisis resources in these simulations, and most were too slow to escalate mental health risk scenarios, postponing referral to a human to potentially dangerous levels.

Hands-On Tutorial

Let's explore how state-of-the-art AI models currently perform as mental health resources for real-world patients. Here is an example of a chatbot that's currently available for anyone to use on the Internet (including your patients) - click on the link to open it in a new window.

Assume the role of a patient seeking mental health support and resources. How accurate is the model as a therapist? How empathetic is the model? Would you use this particular model for mental health support? Why or why not?

Summary

Algorithms are functions that map inputs to outputs. Some algorithms are easy to describe while others are harder to write down. Machine learning is the process of computers learning hard-to-write-down algorithms from past observations, with the goal of learning algorithms that are generalizable to new sets of inputs.

❶ This particular model is hosted on Hugging Face, which has become the *de-facto* website for publishing publicly available machine learning models like the one we're exploring here. Anyone can download a model and use it for applications such as mental health among others.

Additional Readings

1. Topol EJ. High-performance medicine: The convergence of human and artificial intelligence. Nat Med 25: 44-

   56. (2019). doi: 10.1038/s41591-018-0300-7. PMID: 30617339

2. Sidey-Gibbons JAM, Sidey-Gibbons CJ. Machine learning in medicine: A practical introduction. BMC Medical Research Methodology 19(64). (2019). doi: 10.1186/s12874-019-0681-4. PMID: 30890124

3. JAMA Podcast with Dr. Kevin Johnson from Penn Medicine. October 4, 2023. Link.