

# **AI IN HEALTHCARE FOR CLINICIANS**

## **Part 2**

**Abhishek Moturu & Gemma Postill**  
September 2025

# ACKNOWLEDGEMENTS

**This presentation was prepared by:**

- Dr. Laura Rosella
- Dr. Amol Verma
- Dr. Jethro Kwong
- Gemma Postill
- Abhishek Moturu
- Dr. Armaan Malhotra
- Sujay Nagaraj
- Konrad Samsel
- Anglin Dent
- Julie Midroni
- Samantha Unger

# LEARNING OBJECTIVES

1. Define artificial intelligence, machine learning, and deep learning
2. Identify appropriate cautions with potential algorithmic biases
3. Describe potential application of artificial intelligence in clinical practice
4. Practice interpreting and utilizing a deployed model's prediction
5. Recognize medicolegal implications of AI

**Q1**

# Who has heard of AI/ML?

Q2

# Who was here last year?



Temerty Centre for AI Research  
and Education in Medicine

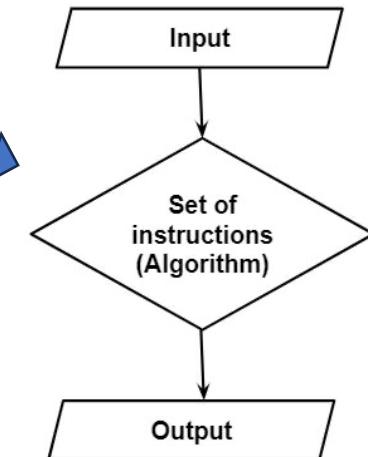
UNIVERSITY OF TORONTO

# 1. INTRO TO AI/ML

The “5-minute version” of AI and  
ML for PGME trainees

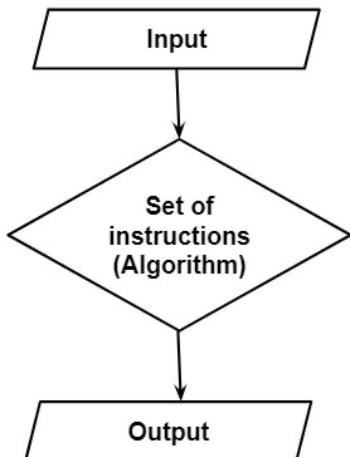
# WHAT IS ARTIFICIAL INTELLIGENCE?

The science of **computer algorithms** able to perform tasks that imitate human **cognitive** functions and intelligence



# WHAT IS ARTIFICIAL INTELLIGENCE?

What does this remind you of?



**APPLE CRUMBLE COFFEE CAKE**

**INGREDIENTS**

- 1 1/2 cups all-purpose flour
- 1/2 cup granulated sugar
- 1/4 cup unsalted butter (melted)
- 1/2 cup milk
- 1 large egg
- 1 teaspoon baking powder
- 1/2 teaspoon ground cinnamon
- 1/4 teaspoon salt
- 2 cups peeled, diced apple
- 1/3 cup all-purpose flour
- 1/3 cup light brown sugar (packed)
- 1/2 teaspoon ground cinnamon

**INSTRUCTIONS**

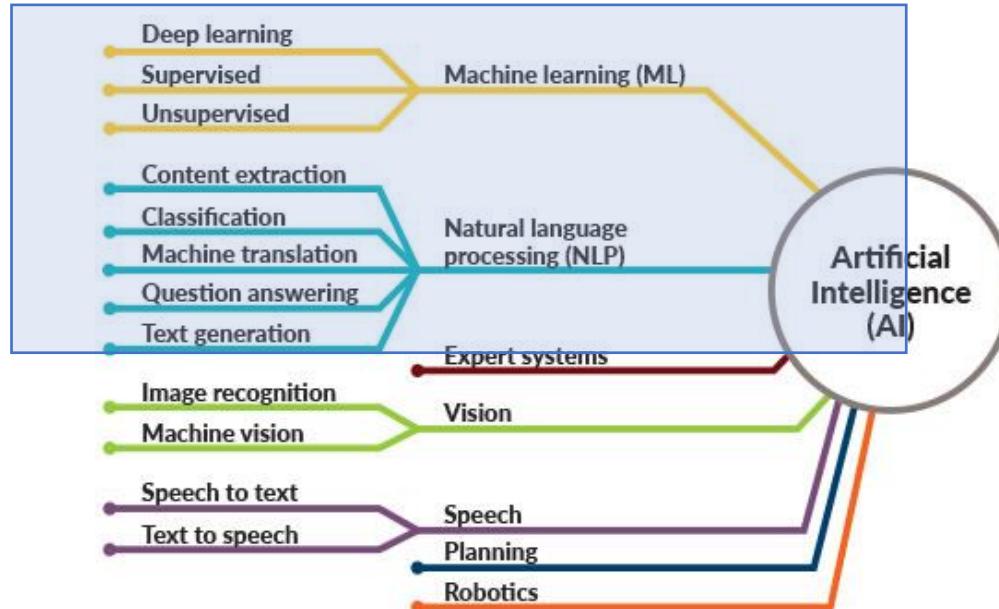
1. Preheat oven to 350°F (175°C) and grease an 8-inch square baking pan.
2. In a large bowl, combine flour, sugar, baking powder, cinnamon, and salt. Add egg, milk, and melted butter; mix until smooth. Fold in apples. Spread in pan.
3. For the crumble topping, mix flour, brown sugar, cinnamon, and softened butter until crumbly.
4. Sprinkle the crumble topping over cake batter. Bake for 35-40 minutes, or until a toothpick inserted in the center comes out clean.

*Recipe generated with ChatGPT*

The image shows a recipe card for "Apple Crumble Coffee Cake". At the top is the title "APPLE CRUMBLE COFFEE CAKE". Below it is a section titled "INGREDIENTS" with a bulleted list of ingredients. To the right of the list is a photograph of a square piece of cake with a crumbly topping. Below the ingredients is a section titled "INSTRUCTIONS" with a numbered list of four steps. At the bottom right is the text "Recipe generated with ChatGPT".



# THERE ARE MANY TYPES OF COMPUTER ALGORITHMS



<https://www.automation.com/en-us/articles/august-2022/ai-machine-learning-human-intelligent-systems>

# WHO INVENTED AI?

1. **Silent Generation** (born between 1925 and 1945)
2. **Boomers** (born between 1946 and 1964)
3. **Gen X** (born between 1965 and 1980)
4. **Millennials** (born between 1981 and 1996)

# BRIEF HISTORY OF AI

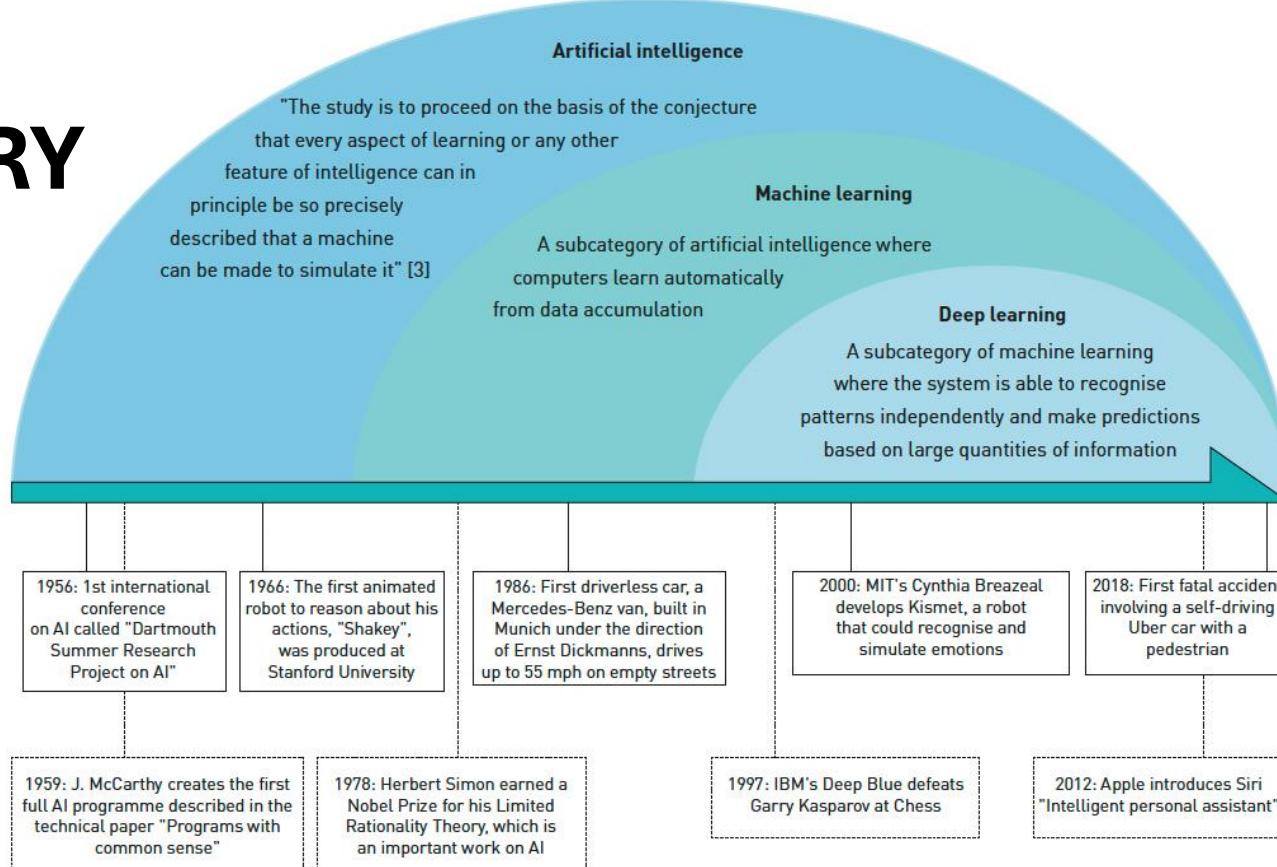


FIGURE 1 Brief history of artificial intelligence (AI).



Temerty Centre for AI Research  
and Education in Medicine

UNIVERSITY OF TORONTO

# WHY IS AI POPULAR NOW?

- Big data
  - Computational power
  - High quality algorithm
- {
- Digital health infrastructure
  - Innovative leadership
  - Ethics and policy
- {
- Technology Advances**
- System Advances**



# **WHY DO WE EVEN NEED AI IN HEALTHCARE?**

**Examples of pressures to innovate in healthcare:**

- Aging populations
- Health workforce shortages
- Rise in chronic disease
- Rural and remote access gaps

# WHY DO YOU NEED TO LEARN ABOUT AI?

*Responsible use of AI is becoming a  
new cornerstone of clinical practice*

## 2. OPPORTUNITIES FOR AI IN MEDICINE

How can ML streamline tasks and provide clinical decision support?



## Emergency Department

Using his  
volumes  
arrive to  
long exper-

## Waiting Room

Waiting  
stressed  
long exper-

Experi-  
Most



## Risk Assess- ment

The Risk  
identifies  
blood su-

There are  
Emergency  
response -  
administra-



## Emergency Department

An interpro-  
fessional  
health discus-  
sion occupa-  
tional absences a-



## Interprofes- sional Team

This AI solu-  
tion can alert  
Departments  
about the patient's  
The predict-  
tive model can  
faster diagno-



## Brain Injury

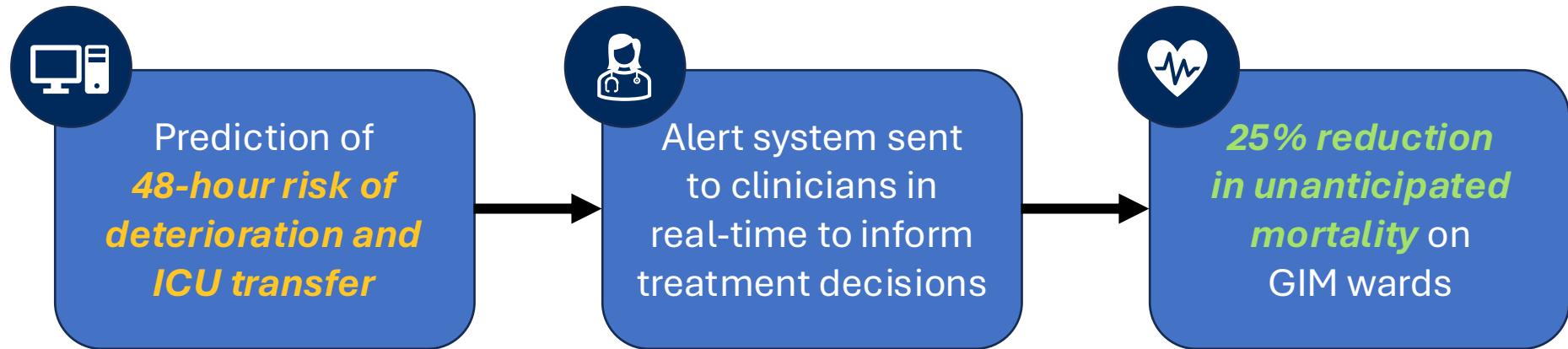
CHARTWatch patient monitoring

CHARTWatch uses patient data from the hospital's existing electronic patient record – such as laboratory test results, patient vitals and patient demographics – to predict the level of support a patient will need.



# HOW CAN AI TRANSFORM PATIENT CARE?

A real-world example from U of T: CHARTwatch, deployed at Unity Health



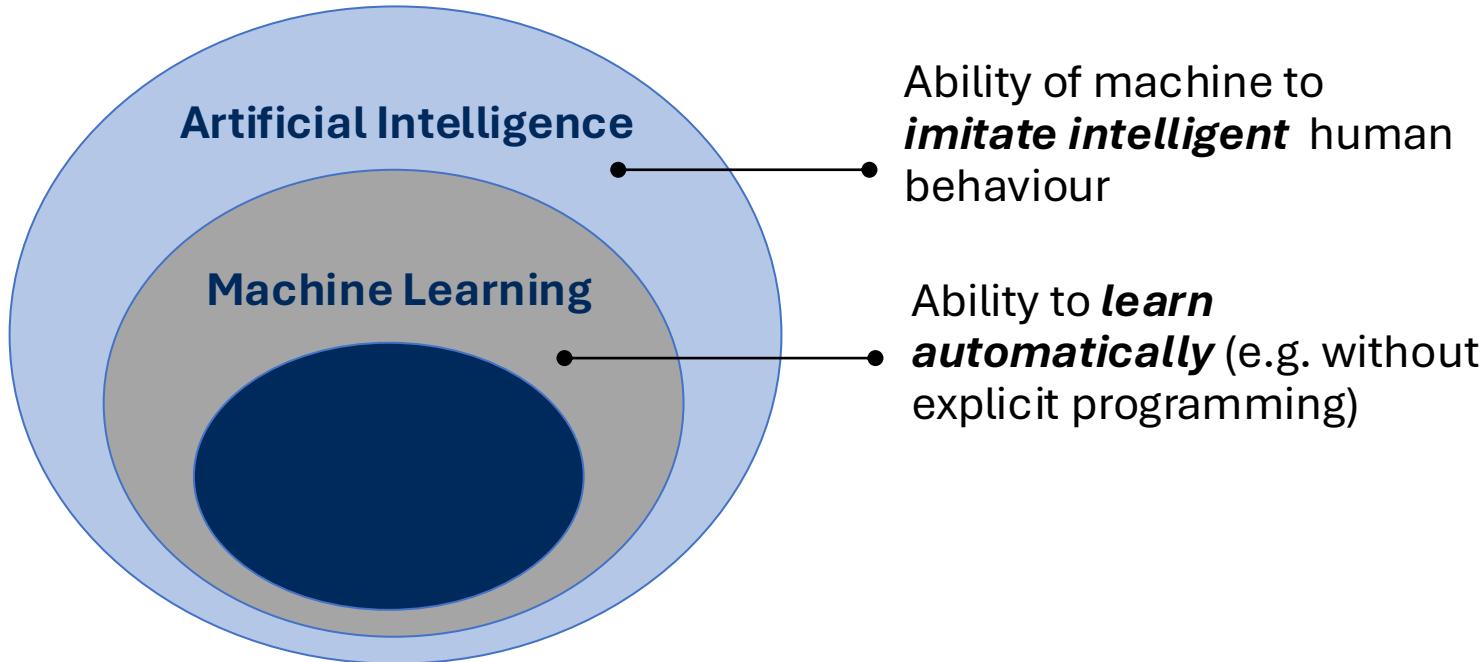
<https://www.cmaj.ca/content/196/30/E1027>

***AI is already being used by clinicians to save lives.***

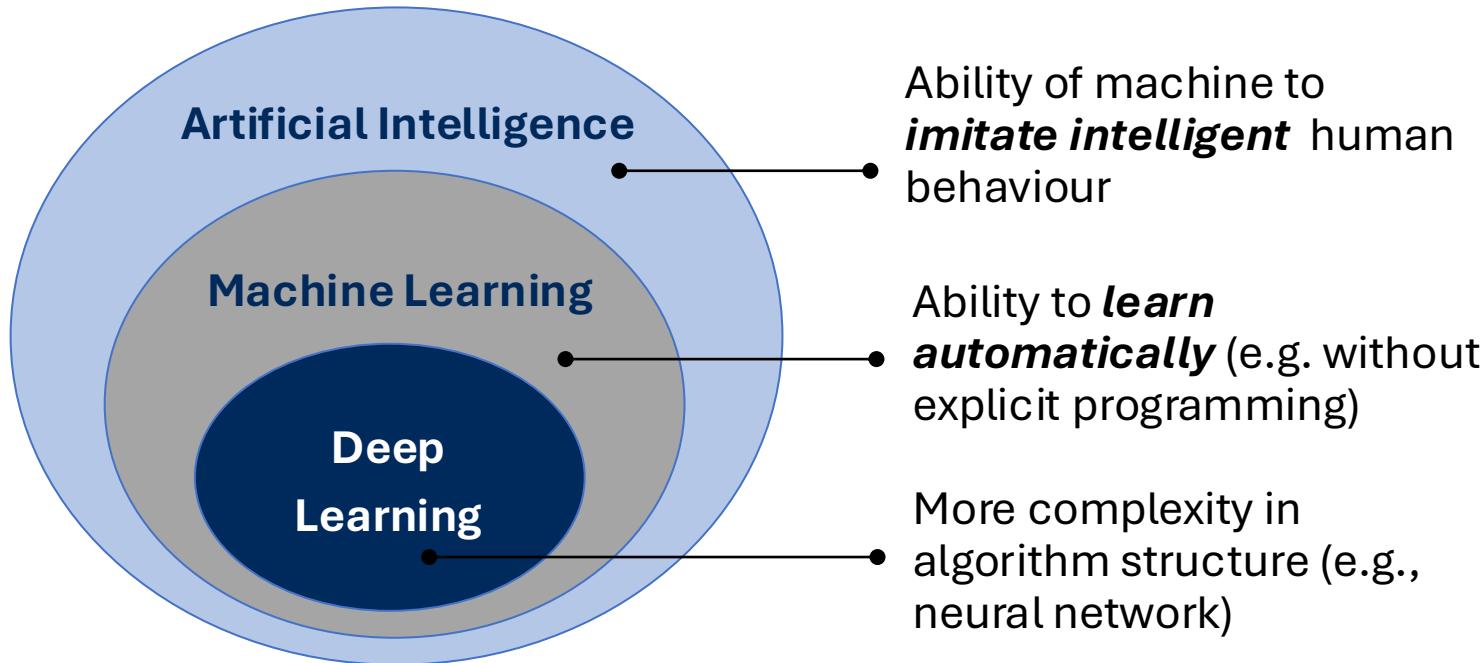
### 3. HOW DOES MACHINE LEARNING WORK?

How does it help with *healthcare* tasks?

# WHAT IS MACHINE LEARNING?



# WHAT IS MACHINE LEARNING?

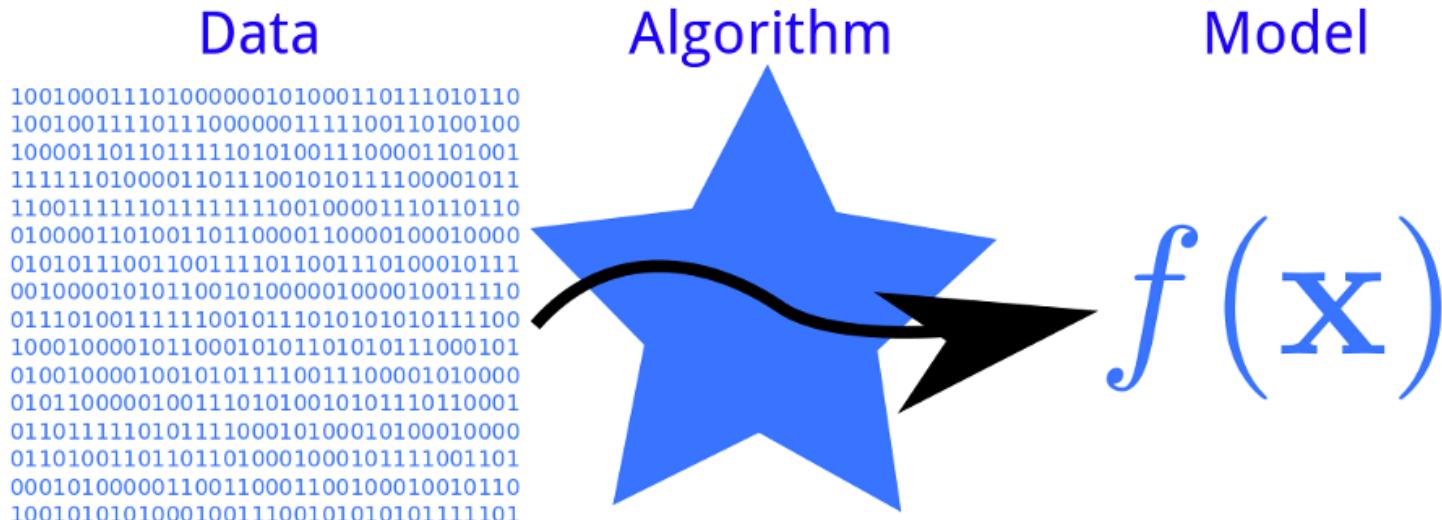


# KEY ML TERMINOLOGY

- **Feature** – akin to a variable in a model
- **Big data** – data of size large enough that traditional manipulation and management is logically/feasibly challenging
- **Algorithm** – a set of rules provided to AI that allows the machine to perform certain tasks, such as classification
- **Model** – an algorithm that has been trained to perform a task

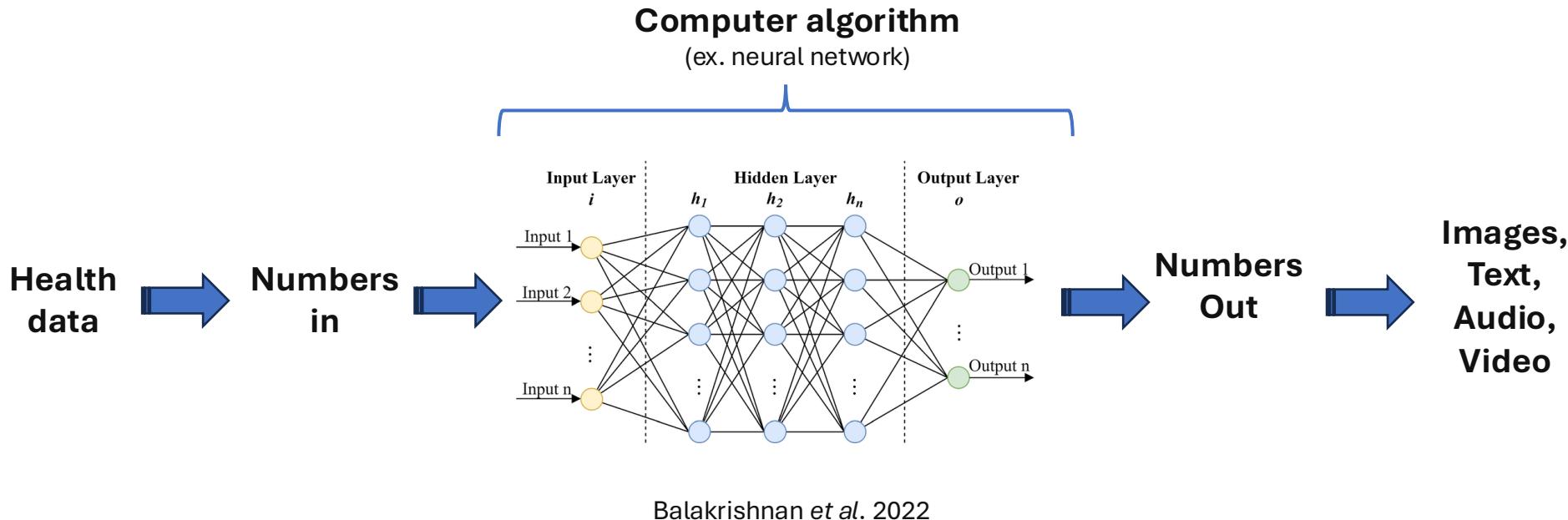
# ML TERMINOLOGY

## ALGORITHM VS. MODEL



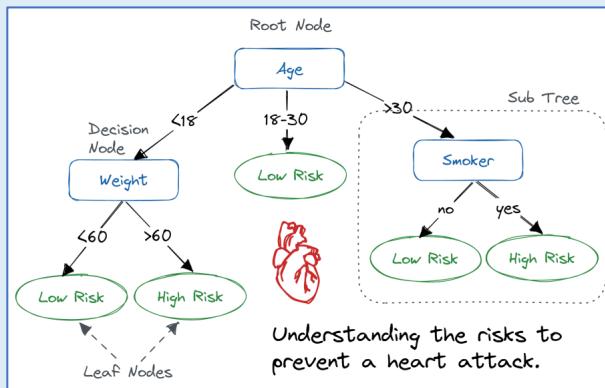
SOURCE: <https://sowjanyasadashiva.medium.com/machine-learning-algorithm-vs-model-2db36e42d127>

# HOW AI/ML WORKS IN GENERAL

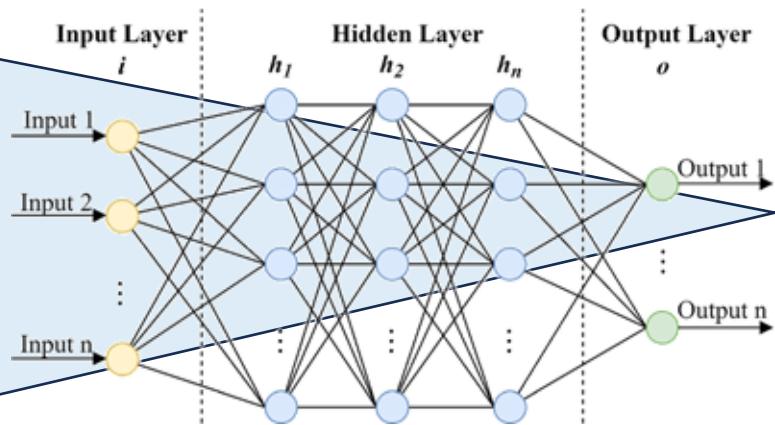


# SPECTRUM OF COMPLEXITY

## Decision Tree



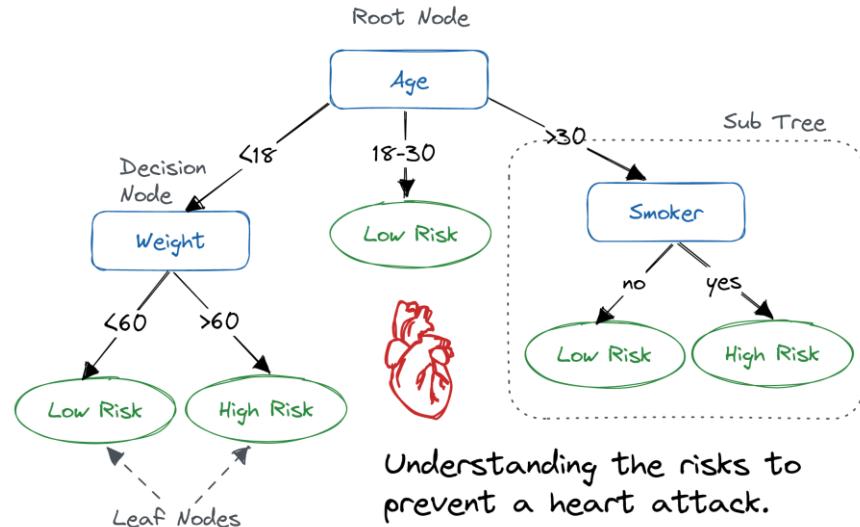
## Neural Network



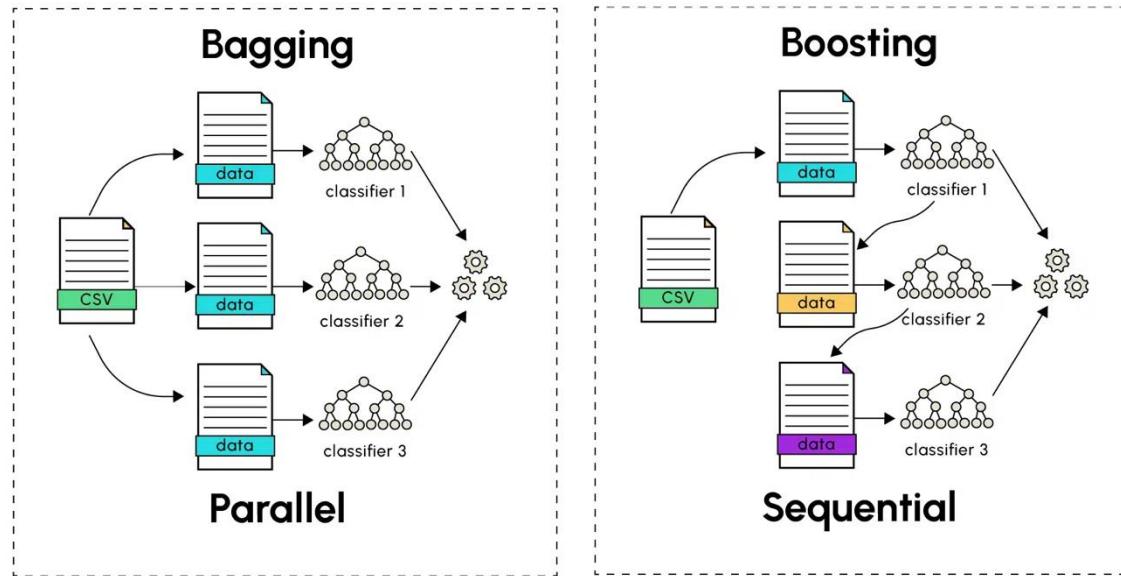
Balakrishnan et al. 2022

# DECISION TREES

- **How it works:** Flowchart-like structure where each question splits the data into smaller groups until a decision is made
- Have been used in healthcare for a long time!
- **When does a decision tree become AI?**



# DECISION TREES – BAGGING AND BOOSTING



# DOES ALL AI WORK THE SAME WAY?

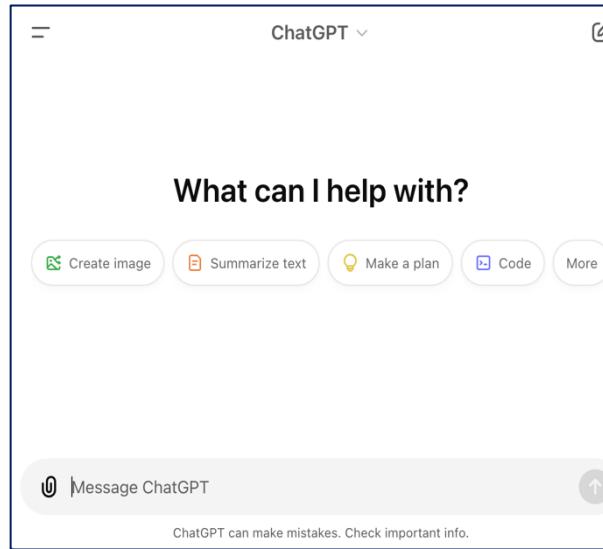
## Non-Generative AI

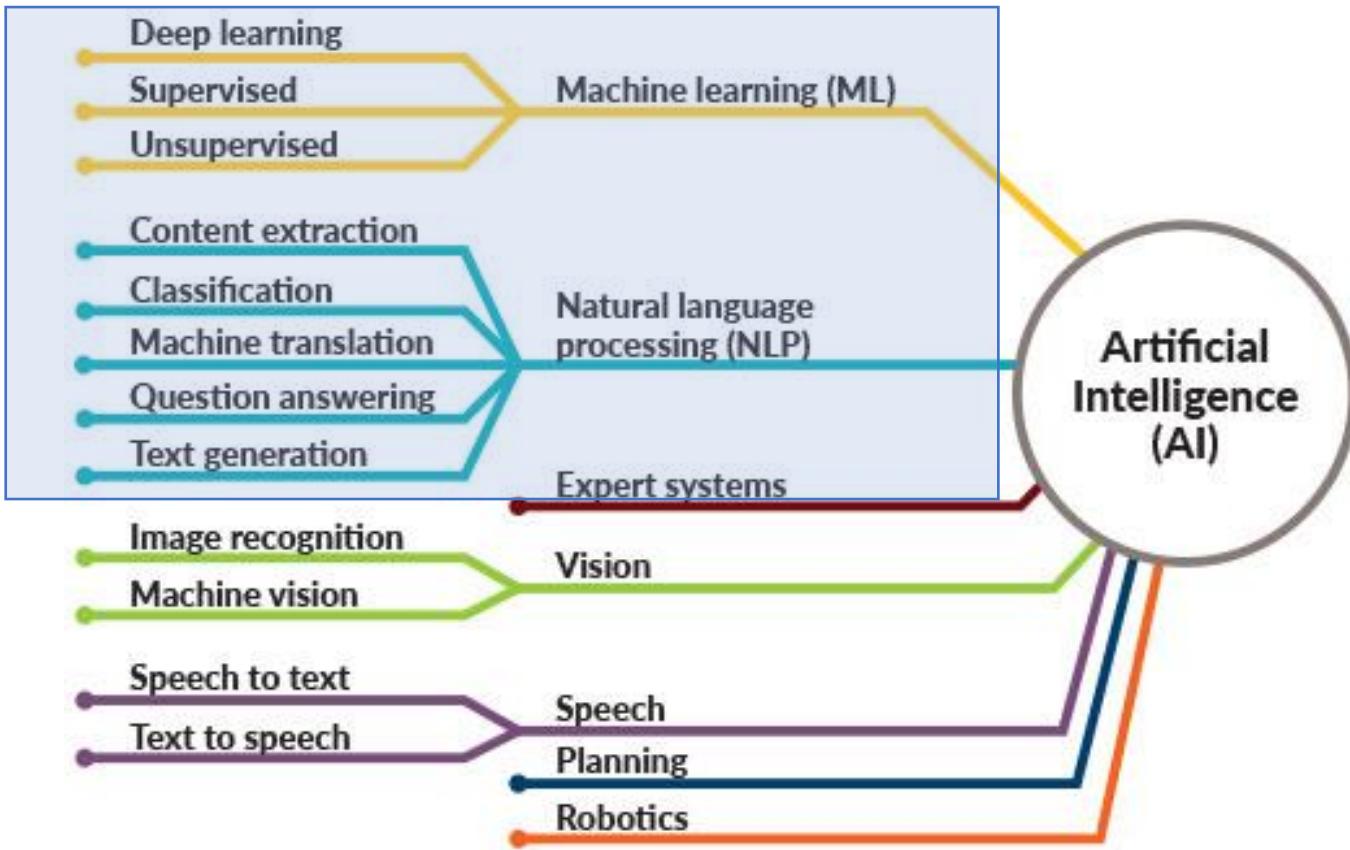


### CHARTWatch patient monitoring

CHARTWatch uses patient data from the hospital's existing electronic patient record – such as laboratory test results, patient vitals and patient demographics – to predict the level of support a patient will need.

## Generative AI





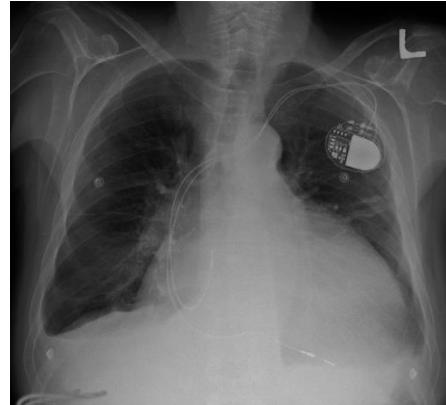
# RECAP OF NON-GENERATIVE ML

Using an example focused on imaged-based  
**supervised machine learning**

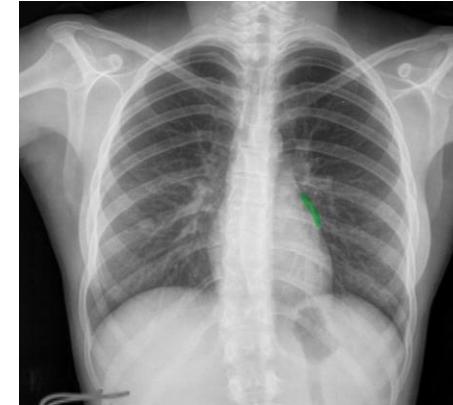
# SUPERVISED LEARNING EXAMPLE



Cardiomegaly



Cardiomegaly



Normal



**ML Task:**

Model learns  
patterns between  
image and label of  
cardiomegaly

# CASE STUDY OF AI FOR A CLINICAL PROBLEM

An Early Warning System for the Real-time Image-Based Triage of Patients Suffering Traumatic Brain Injury:  
**Development of ASIST-TBI**

Slides developed by Dr. Armaan Malhotra

# Progress Today

Original Research

## Vision Transformer-based Decision Support for Neurosurgical Intervention in Acute Traumatic Brain Injury: Automated Surgical Intervention Support Tool (ASIST-TBI)

Christopher W. Smith\*, Armaan K. Malhotra\*, Christopher Hammill, Derek Beaton, Erin M. Harrington, Yingshi He, Husain Shakil, Amanda McFarlan, Blair Jonesn, Hui Ming Lin, François Mathieu, Avery B. Nathens, Alun D. Ackery, Garrick Mok, ... See all authors ▾

\* C.W.S. and A.K.M. contributed equally to this work.

\*\* E.C. and C.D.W. are co-senior authors.

Author Affiliations

Published Online: Jan 10 2024 | <https://doi.org/10.1148/ryai.230088>



## 4. HOW DOES NATURAL LANGUAGE PROCESSING WORK

Focusing on the example of **Large Language  
Models (LLMs)**

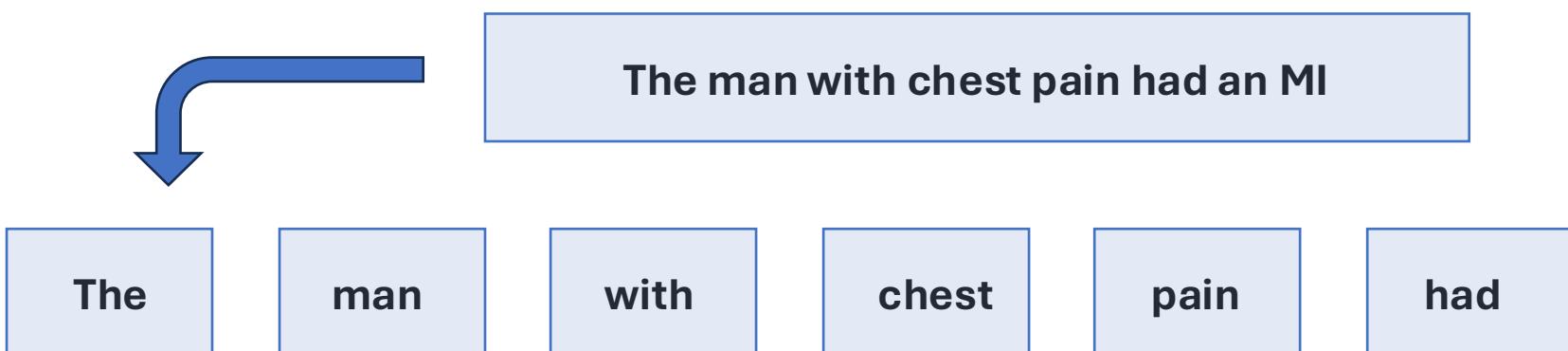


# LARGE LANGUAGE MODELS

- LLMs = Large Language Model
- Understand and generate human-like text
- Use a probability function to tokenize and predict string (= words / documents)

# HOW LLMs WORK: TOKENS

Language is broken down and manipulated mathematically by LLMs into tokens, which are then represented numerically in models



# HOW LLMs WORK: PROBABILITIES

## Training Data

The man with chest pain had an MI

The man with chest pain had an MI

The man with chest pain had an MI

The man with chest pain had an anxiety

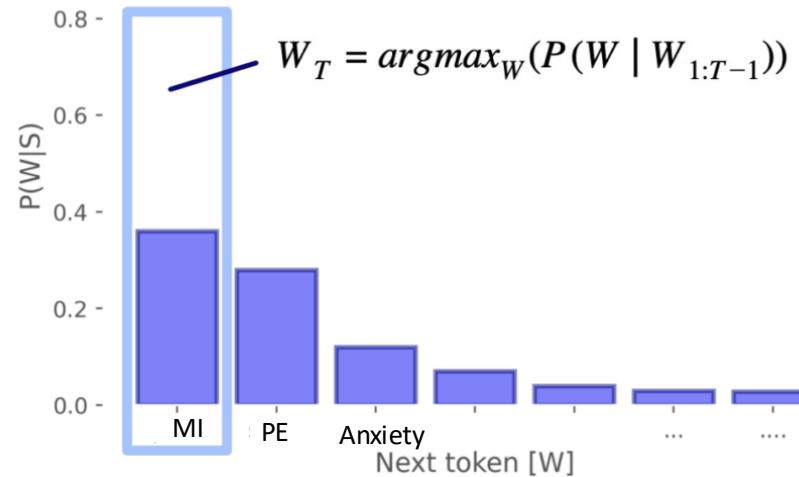
The man with chest pain had a PE

The man with chest pain had an anxiety

The man with chest pain had a PE

The man with chest pain had an MI

The man with chest pain had \_\_



# HOW LLMs WORK: ATTENTION

What word is the most important in predicting the clinical condition (next word)?

The

man

with

chest

pain

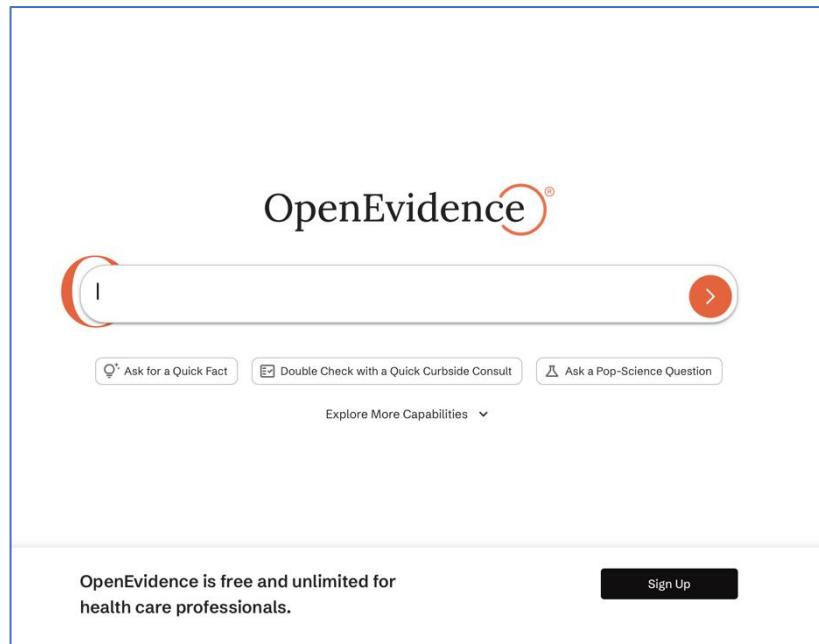
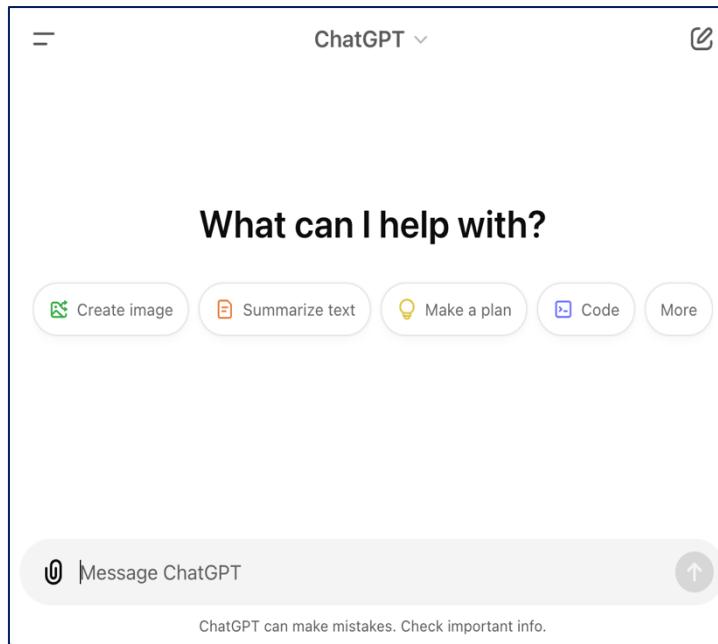
had



***LLMs were revolutionized by the concept of “attention”***

Not the most recent word necessarily!

# THE RESULT OF ATTENTION + BIG DATA



# TYPES OF INCORRECT LLM OUTPUTS

Category	Example
<b>Myths</b> – Errors in training data	Ex. An apple a day keeps the <b>doctor</b> away
<b>Mistakes</b> – Violates rules	Ex. Joe was diagnosed in 2010. In 2020, he has had the condition for <b>6 years</b> .
<b>Hallucinations</b> – false or misleading information presented as facts (plausible but no clear origin)	Ex. Write a consult for James... <i>see next slide</i>

# HALLUCINATIONS

*Can you spot the hallucination?*

Write a consult from a cardiac surgeon for James who has had an MI



Here's a draft for the consult note:

Cardiac Surgery Consult

Answer in chat instead >

Does this outline address all the necessary details for James's case? Let me know if you'd like to tailor any part further!



## History of Present Illness:

James is a [age]-year-old male with a significant past medical history of [list pertinent history: hypertension, hyperlipidemia, diabetes, etc.], who presented to [hospital name] on [date] with symptoms of acute chest pain radiating to the left arm and jaw, associated with diaphoresis and shortness of breath. He was diagnosed with an anterior ST-elevation myocardial infarction (STEMI) via ECG and elevated cardiac biomarkers (troponin [insert values]).

He underwent emergent [PCI (percutaneous coronary intervention) with stent placement] to the [specific vessel, e.g., LAD - left anterior descending artery] with partial resolution of symptoms. However, echocardiography performed post-intervention revealed:

- LVEF: [e.g., 35%]
- Regional wall motion abnormalities: [e.g., anterior and apical hypokinesis]
- [Optional findings, e.g., moderate MR (mitral regurgitation) or ventricular septal defect].

Cardiology now requests a surgical evaluation for potential [CABG (coronary artery bypass grafting), VSD repair, mitral valve repair, etc.].



Temerty Centre for AI Research  
and Education in Medicine

UNIVERSITY OF TORONTO

# HALLUCINATIONS – WHY?

- Bad training data (ex. false information)
- Tricky prompts
- “Unlearnable” fact distributions (ex. rare fact rates)
  - If information occurs once, perceived as random
  - Proportion of times “rare facts” occur is related to the proportion of hallucinations

# HALLUCINATIONS – Prompting

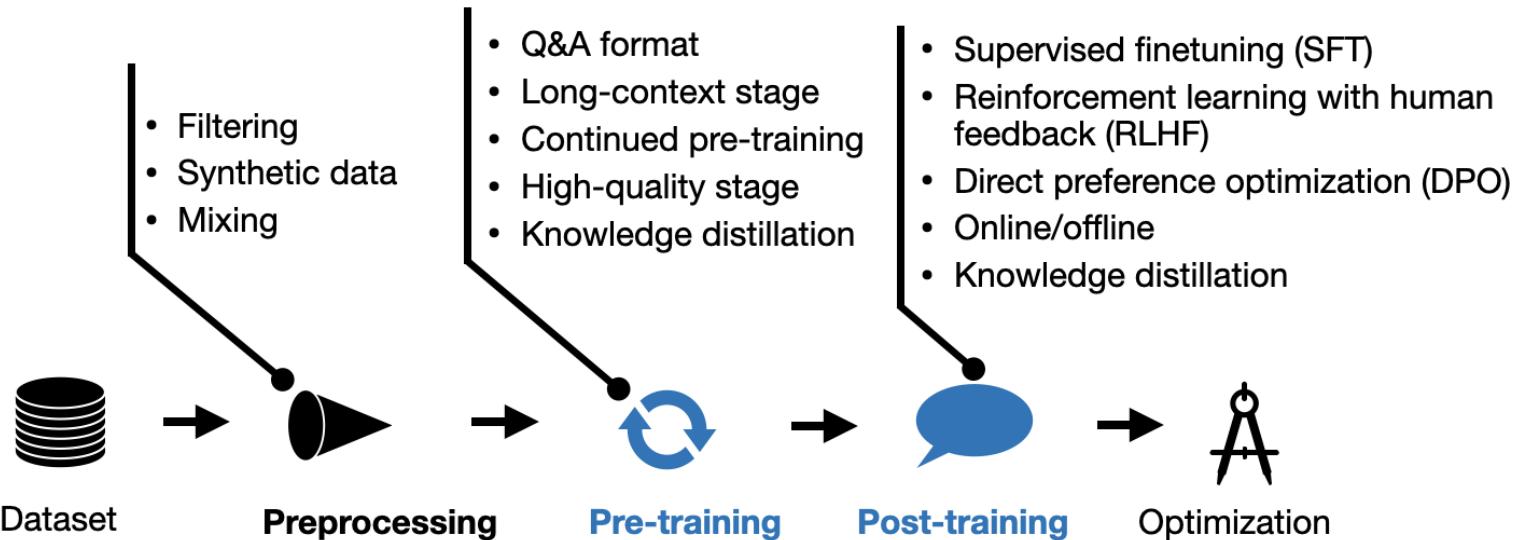
## Improve prompts by :

- Being clear and specific
- Providing relevant context and background information
- Providing examples
- Breaking down complex tasks
- Tell LLM to slow down or to confirm

## Use the 3xPs:

- Persona (*you are a cardiac surgeon*)
- Purpose (*creating a clinical note*)
- Parameters (*you only know...*)

# HALLUCINATIONS: Reduced with Post-Training



Source: <https://magazine.sebastianraschka.com/p/new-llm-pre-training-and-post-training>

## Summary

# DOES ALL AI WORK THE SAME WAY?

### Non-Generative AI

- Predicts output/label/answer based on learned relationships
- Concerns for quality of output if data (person/patient) has not been in the training data

### Generative AI

- Generates new content
- Concerns for hallucinations

*Note: models with image output can also be generative*

## 5. CAUTION WITH AI

How do you know if you have a good AI?

## Word2Vec Model

- > Paris – France + Italy  
Rome
- > King – Man + Woman  
Queen

# Word2Vec Model

- > Doctor – Man + Woman  
Nurse

“We dug deeper and saw that it was even worse than that”

- > Computer Programmer – Man + Woman  
Homemaker

“We jointly realized – Hey, there’s something wrong here”

Bolukbasi et al. (2016) <https://arxiv.org/abs/1607.06520>

# ML CAN BE BIASED!

**Inadequate label:** healthcare costs as a measure of healthcare needs

Race was not included yet the model was biased!

The screenshot shows a web page from the journal **Science**. At the top, there are navigation links: **Current Issue**, **First release papers**, **Archive**, **About** (with a dropdown arrow), and a button labeled **Submit manuscript**. Below the header, the URL indicates the article is in **SCIENCE**, volume 366, number 6464. The main title of the article is **Dissecting racial bias in an algorithm used to manage the health of populations**. Below the title, the authors listed are **ZIAD OBERMEYER**, **BRIAN POWERS**, **CHRISTINE VOGELI**, and **SENDHIL MULLAINATHAN**. There is also a link to **Authors Info & Affiliations**. Social media sharing icons for Facebook, Twitter, LinkedIn, and others are visible on the right side of the article summary.

# **SELECT CAUTIONS**

1. Patient Population
2. Label Quality
3. Overfitting
4. Interpretability
5. Dataset Shift

## 2. LABEL QUALITY

Imagine your model is predicting hypertension

**What is the best label for hypertension?**



## 2. LABEL QUALITY

What is the best label for hypertension?

How should we measure it?

Sphygmomanometer



Manual



Semi-Automatic



Automatic

<https://alder.in/?a=181917915>

## 2. LABEL QUALITY

**What is the best label for hypertension?**

**When should we measure it?**



<https://alder.in/?a=181917915>

## 2. LABEL QUALITY

What is the best label for **hypertension?**

Why should we measure it?

Treatment!

Treatment thresholds by age and chronic condition

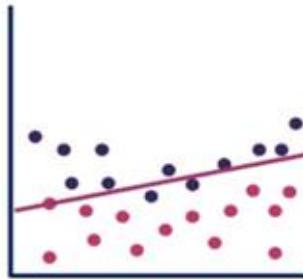
Age group	Office SBP treatment threshold (mmHg)					Office DBP treatment threshold (mmHg)
	Hypertension	+ Diabetes	+ CKD	+ CAD	+ Stroke/TIA	
18 - 65 years	≥140	≥140	≥140	≥140 <sup>a</sup>	≥140 <sup>a</sup>	≥90
65 - 79 years	≥140	≥140	≥140	≥140 <sup>a</sup>	≥140 <sup>a</sup>	≥90
≥80 years	≥160	≥160	≥160	≥160	≥160	≥90
Office DBP treatment threshold (mmHg)	≥90	≥90	≥90	≥90	≥90	

BP = blood pressure; CAD = coronary artery disease; CKD = chronic kidney disease; DBP = diastolic blood pressure; SBP = systolic blood pressure; TIA = transient ischaemic attack.

<sup>a</sup>Treatment may be considered in these very high-risk patients with high-normal SBP (i.e. SBP 130–140 mmHg).

### 3. MODEL FIT

*Imagine the goal of this algorithm is to differentiate between purple & pink*



**Underfitting**

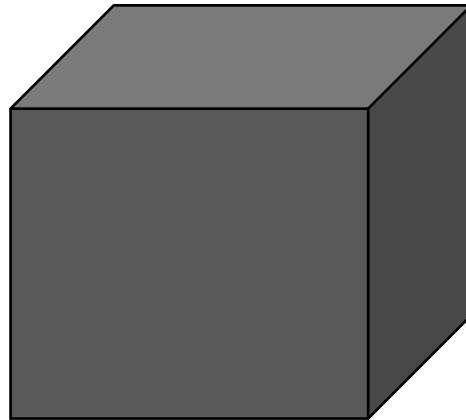


Poor performance  
on testing and  
training data

Source: <https://www.r-bloggers.com/2022/09/how-to-avoid-overfitting/>

# 4. MODEL EXPLAINABILITY

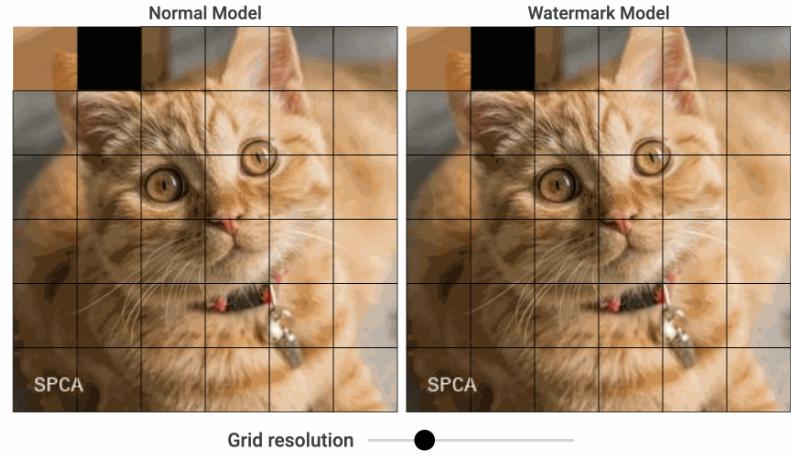
**Isn't it all  
black box AI?**



# 4. MODEL EXPLAINABILITY

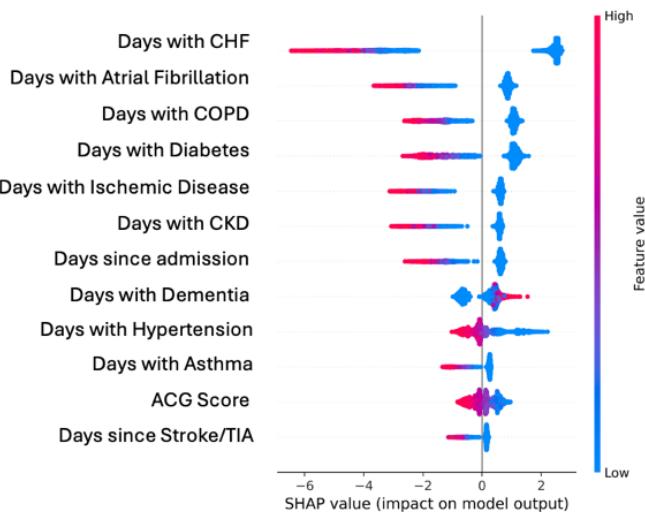
**Explainability:** Justification  
for model outputs

- Especially useful when dealing with less interpretable models
- Model explanations: SHAPley, LIME, Grad-CAM, CAVs

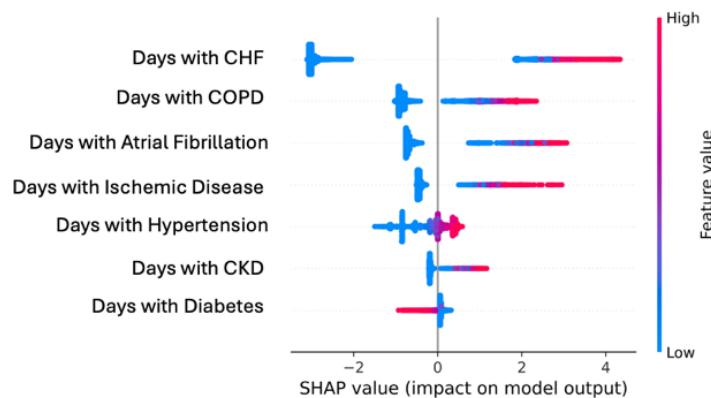


# 4. MODEL EXPLAINABILITY

**(A) Feature importance of frail model – Cluster B**  
Measured as utility of features (and their values) to predict Cluster B (vs. Cluster A)



**(B) Feature importance of non-frail model – Cluster C**  
Measured as utility of features (and their values) to predict Cluster C (vs. Cluster D or E)



# 5. DATASET SHIFT

**What can change in data that is input into models?**

- New event / diseases (e.g., COVID)
- Changes in disease classification and procedures for treatment
- Changes to data collection process (e.g., automated vitals recording vs. prn)

# **DOES ALL AI WORK THE SAME WAY?**

## **Fixed Models**

- Model weights are frozen
- Applies learned relationships
- Concerns for dataset shift
- Performance can improve (or degrade) overtime

## **Continuous Learning Models**

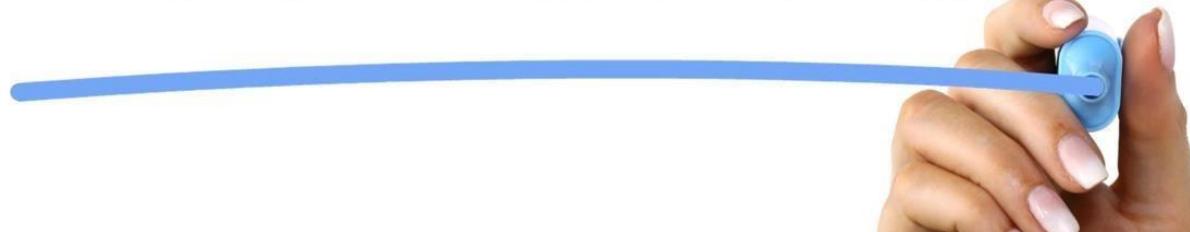
- Model weights are not fixed
- Applies learned relationships
- Model continues to learn when deployed
- Performance can improve (or degrade) overtime

# 6. LIABILITY WITH AI

What responsibility do I have using an AI model?

# LIABILITY OF PHYSICIANS

STANDARD  
OF CARE



# LIABILITY WITH AI

## AI Model for Acute Appendicitis Management

### Why?

- Reduce unnecessary operations
- Reduce radiation exposure
- Improve time to treatment
- Overall patient flow

Data from POCUS  
+ triage vitals

Predictive AI model produces management recommendation

U/I allow automated ordering of antibiotics +/- CT (management indicated)

## Model Interactive User Interface on EMR

### Management of Appendicitis

Assistive AI for use with clinician oversight



#### Vitals

Heart rate: 115 bpm  
Blood pressure: 125/86 mmHg  
Temperature: 38,3°C

#### Recommendation

Recommend operative management

Operative

Non-operative

Image generated with ChatGPT



Temerty Centre for AI Research  
and Education in Medicine

UNIVERSITY OF TORONTO

# **What is the standard of care for managing acute appendicitis?**

# LIABILITY WITH AI

**Case 1. A 25-year-old HR 84, BP 122/74, T 37.7°C. Bedside US: non-perforated, compressible appendix with wall thickening; no abscess; well-appearing.**

- **AI Recommendation:** Non-operative (IV → PO antibiotics, close follow-up)
- **Clinical team's action:** Agree with AI
- **Patient's outcome:** Symptoms resolve, no recurrence at 1-year

**Does physician  
agree with AI?**



**Standard of  
care provided?**



# LIABILITY WITH AI

**Case 2. A 52-year-old, HR 118, BP 98/62, T 38.9°C, peritoneal signs. US: free fluid, non-visualized tip, dirty shadowing.**

- **AI Recommendation:** Operative (urgent lap appendectomy + source control)
- **Clinical team's action:** Surgeon rejects, tries antibiotics-only.
- **Patient's outcome:** Patient deteriorates, septic shock, emergent surgery next day.

**Does physician agree with AI?**



**Standard of care provided?**



**Is physician liable for the bad outcome?**



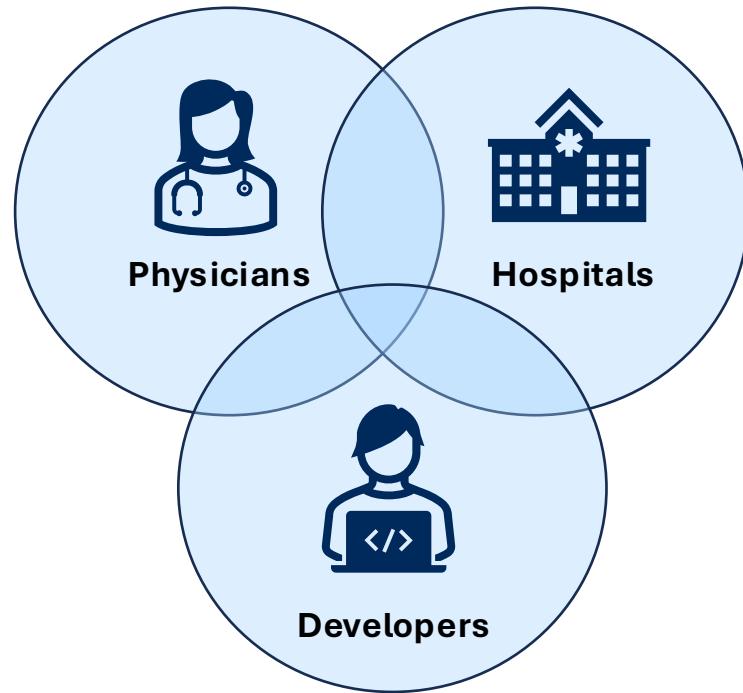
# LIABILITY WITH AI

Figure. Examples of Potential Legal Outcomes Related to AI Use in Clinical Practice

Scenario	AI recommendation	AI accuracy	Physician action	Patient outcome	Legal outcome (probable)
1	Standard of care	Correct	Follows	Good	No injury and no liability
2			Rejects	Bad	Injury and liability
3		Incorrect (standard of care is incorrect)	Follows	Bad	Injury but no liability
4			Rejects	Good	No injury and no liability
5	Nonstandard care	Correct (standard of care is incorrect)	Follows	Good	No injury and no liability
6			Rejects	Bad	Injury but no liability
7		Incorrect	Follows	Bad	Injury and liability
8			Rejects	Good	No injury and no liability

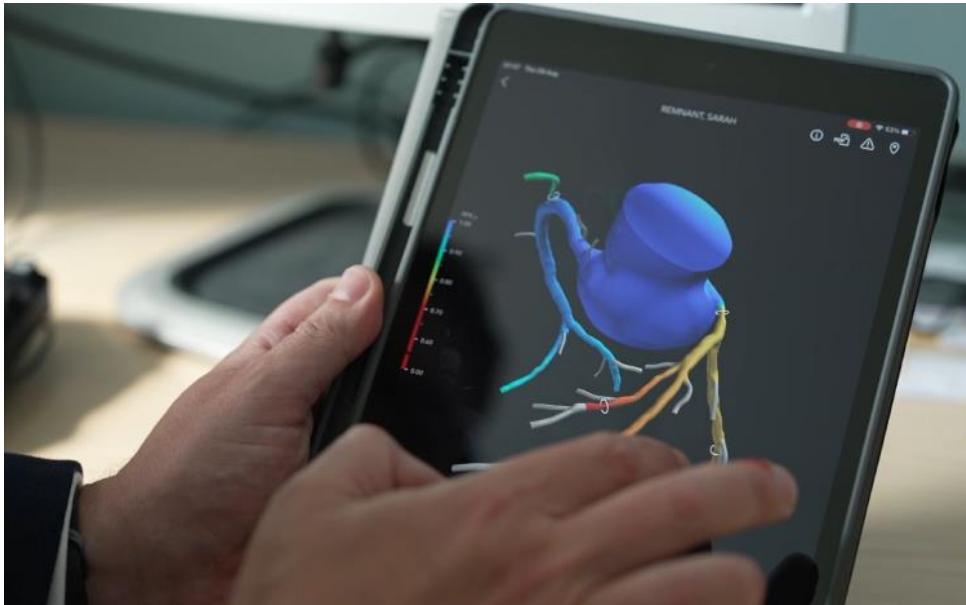
<https://jamanetwork.com/journals/jama/article-abstract/2752750>

# LIABILITY WITH AI



As the quality of AI improves, this is shifting...

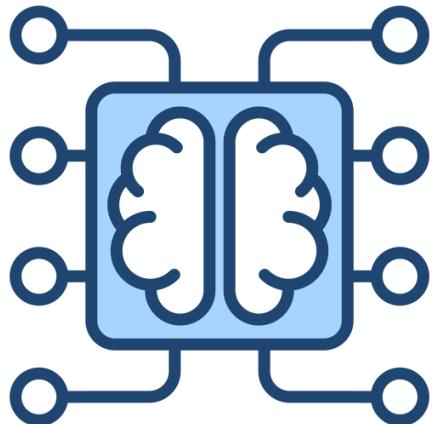
# AI AS STANDARD OF CARE



**HeartFlow** has been rolled out across 56 UK hospitals. Reduced 16% of angiograms

Fairbairn et al. 2025 *Nature medicine*.  
<https://doi.org/10.1038/s41591-025-03620-y>

# HUMAN + AI COLLABORATION



**Is this the best  
combination?**

## 7. INCREASING YOUR AI LITERACY TODAY

Simple ways to follow and engage with emerging AI information

# T-CAIREM Residency AI Programming

**Part 1.**  
Foundations of  
AI

**Part 2.** Day-to-  
day use of AI

**Part 3.** Machine  
Learning for  
Clinical Decision  
Support

**Part 4.** Large  
Language  
Models in  
Clinical Practice

**Part 5.** Deep  
Learning in  
Medical Imaging

**Part 6.** Models  
and Evaluation  
Metrics

**Part 7.** Ethical  
Considerations  
for AI in  
Healthcare



# T-CAIREM Residency AI Bootcamp



**Detail to come...**

# Keeping Up with AI

## NEJM AI Grand Rounds

NEJM AI Grand Rounds, hosted by Arjun (Raj) Manrai, Ph.D. and Andrew Beam, Ph.D., features informal conversations with a variety of unique experts exploring the deep issues at the intersection of artificial intelligence, machine learning, and medicine. You'll learn how AI will change clinical practice and healthcare, how it will impact the patient experience, and about the people who are pushing for innovation. Whether you are an AI researcher or a practicing clinician, these conversations will enlighten and surprise you as we journey through this very exciting field. Produced by NEJM Group.



**Podcasts:** NEJM AI Grand Rounds podcast features informal conversations between experts.

**AI-Specific Subspecialty Journals:** NEJM AI, Nature Digital Medicine, Lancet Digital Health, Radiology: AI, etc.

**T-CAIREM:** Sign up for newsletter/email to get information about educational rounds, events, talks, and conferences.

**Use it:** Everyone is busy, but in this period of rapidly changing technologies, the best way to keep up is to use emerging solutions in real-time so you can form your own opinions.



Temerty Centre for AI Research  
and Education in Medicine

UNIVERSITY OF TORONTO

# THANK YOU!

## Any questions?

[gemma.postill@utoronto.ca](mailto:gemma.postill@utoronto.ca)

[abhishek.moturu@mail.utoronto.ca](mailto:abhishek.moturu@mail.utoronto.ca)

# **TYPES OF MACHINE LEARNING**

Is all machine learning the same? What if the label quality isn't good?

# WHAT IS MACHINE LEARNING?



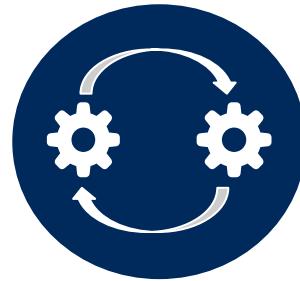
## 1. Supervised Learning

Training a computer model to understand the **relationship between data and its label** (e.g., outcome)



## 2. Unsupervised Learning

Training a computer model to understand **patterns within data without explicit guidance of the label** (e.g., outcome)

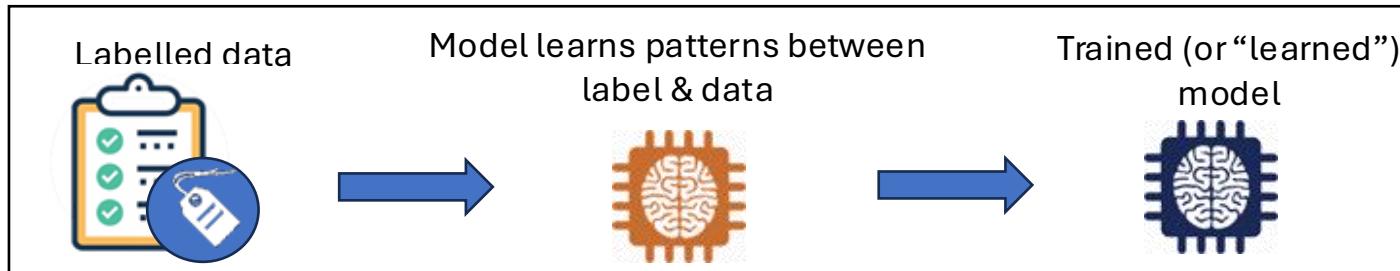


## 3. Reinforcement Learning

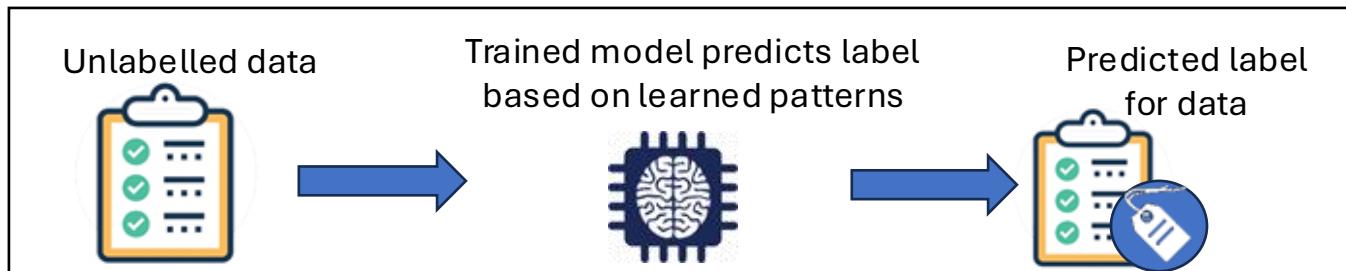
Train a computer model through iterative feedback (trial and error) so that it may **optimize task performance** over time

# SUPERVISED LEARNING

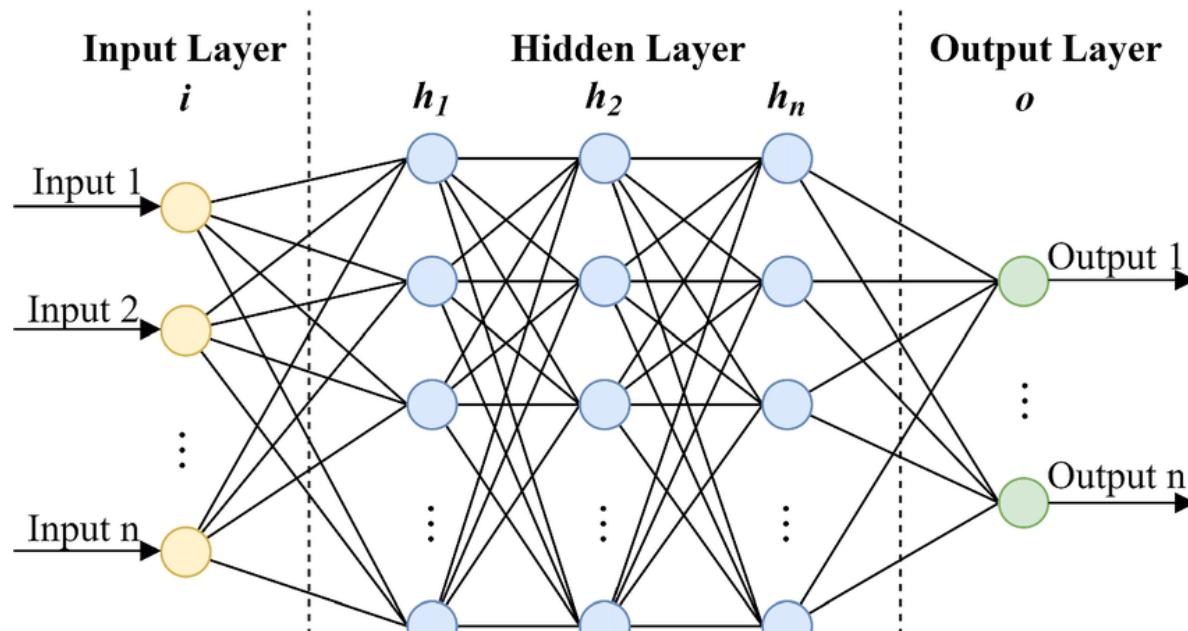
## Developing a Supervised Model (“Model Training”)



## Using a Supervised Model



# HOW AI WORKS – MODEL STRUCTURE

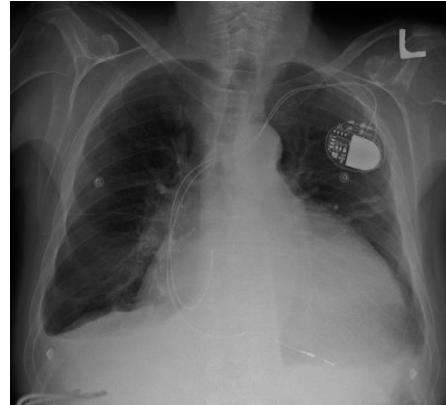


Balakrishnan et al. 2022

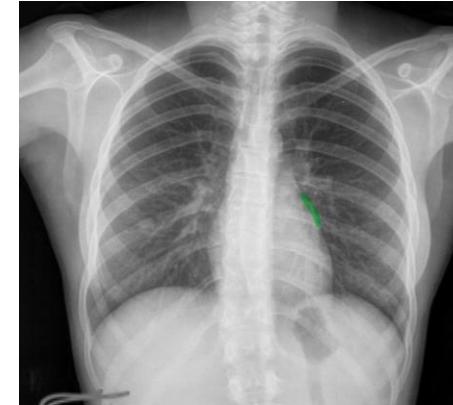
# SUPERVISED LEARNING EXAMPLE



Cardiomegaly



Cardiomegaly



Normal

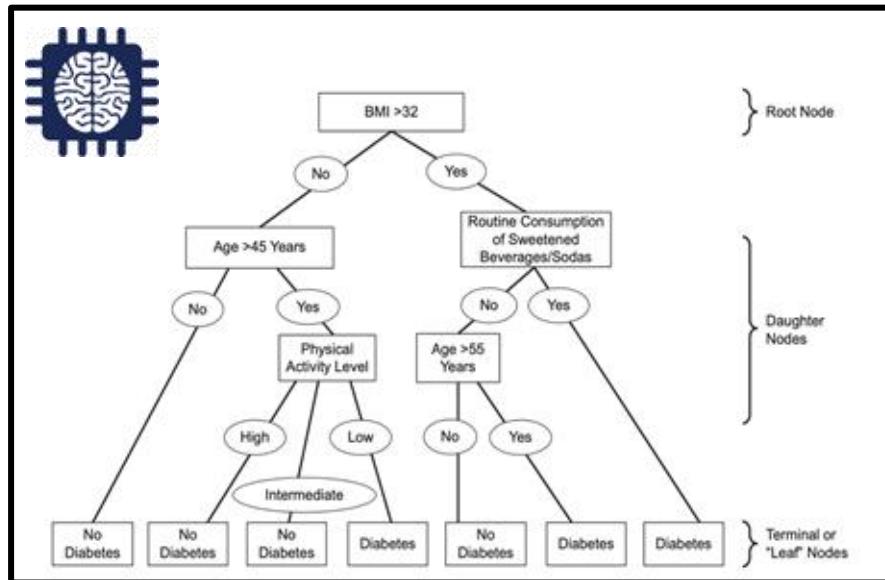


**ML Task:**

Model learns  
patterns between  
image and label of  
cardiomegaly

# SUPERVISED LEARNING EXAMPLE

**ML Task:**  
From patient data,  
predict whether  
patient has diabetes?



*Bi et al. 19, AJE*

**Algorithm: Random Forest**

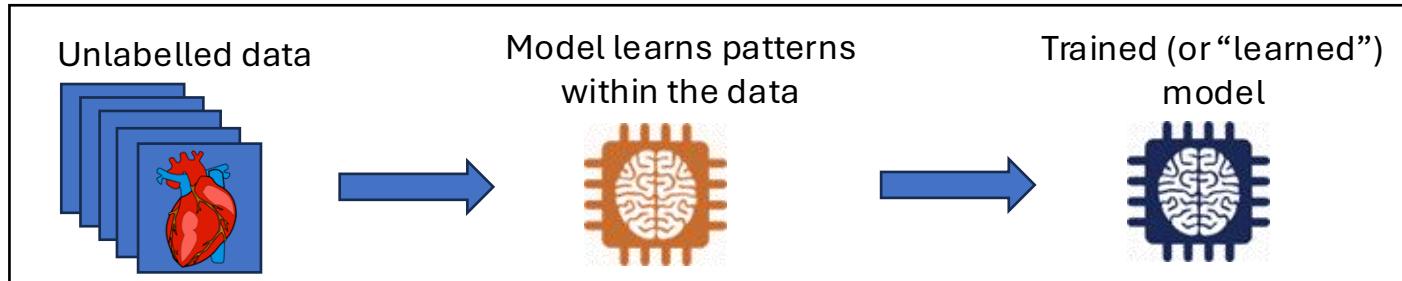


Temerty Centre for AI Research  
and Education in Medicine

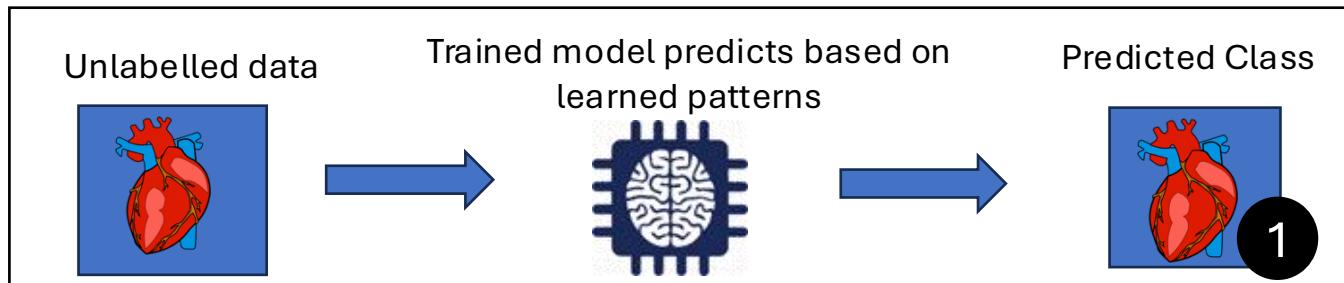
UNIVERSITY OF TORONTO

# UNSUPERVISED LEARNING

## Developing an Unsupervised Model (“Model Training”)



## Using an Unsupervised Model

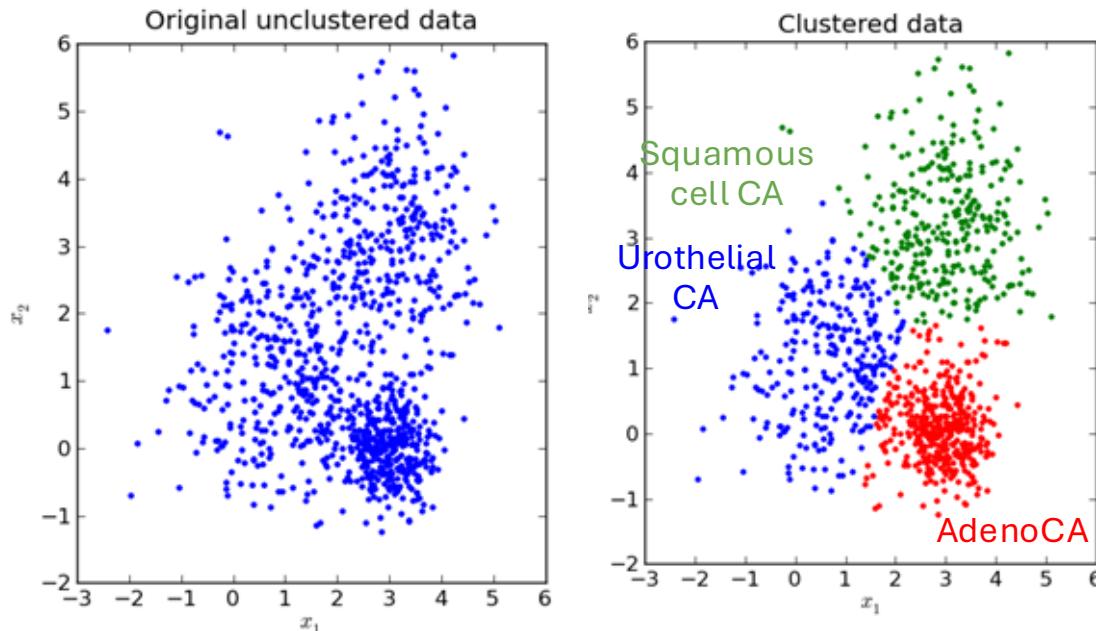


# UNSUPERVISED LEARNING EXAMPLE

**Hypothetical Example:**

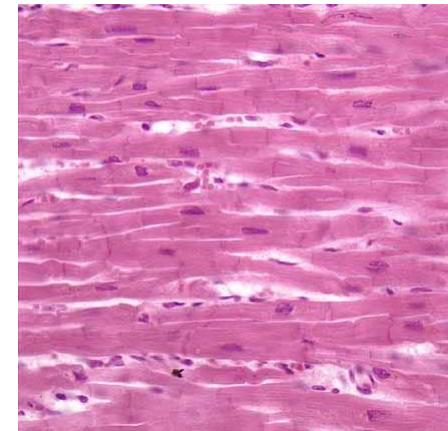
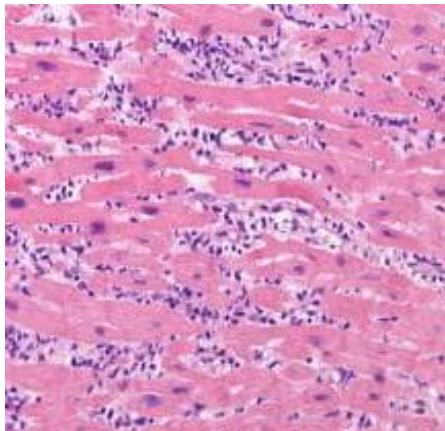
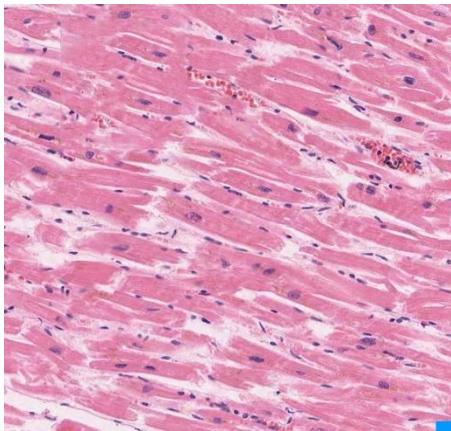
**DATASET** = tumor pathology characteristics of bladder cancer

**GOAL OF MODEL** = distinguish identify cancer subtypes from image characteristics



# UNSUPERVISED LEARNING EXAMPLE

## Cardiac Muscle Tissue Slides



**ML Task:**  
Create risk categories for  
transplant rejection using  
post-transplant histology



Temerty Centre for AI Research  
and Education in Medicine

UNIVERSITY OF TORONTO

# APPRAISING AI TOOLS

How can I tell if the results from AI are good?

**Q3**

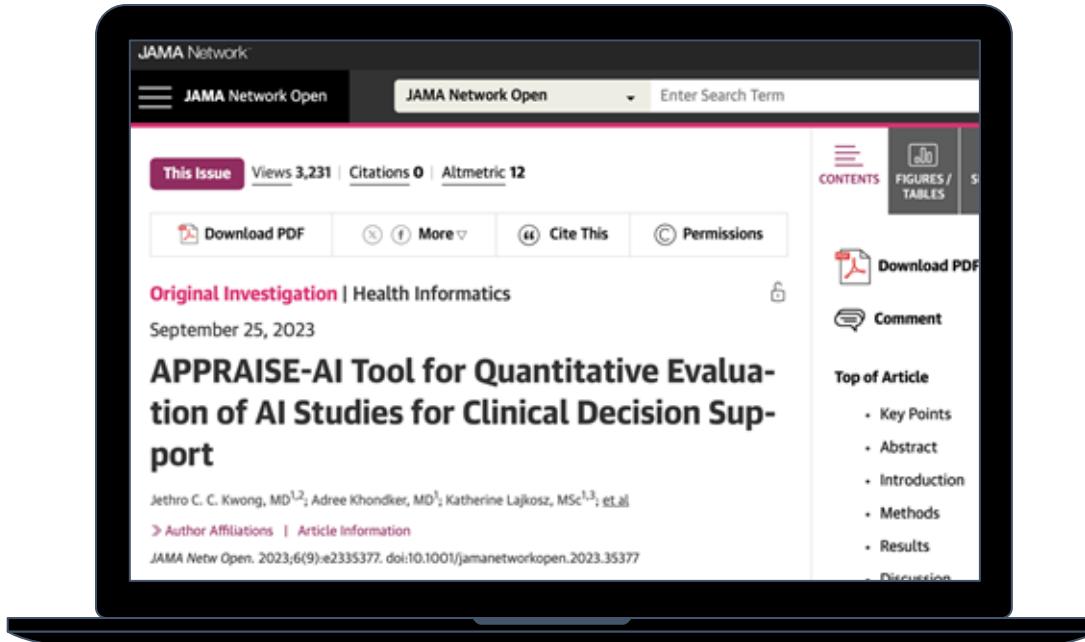
# **Are you confident in appraising AI studies and tools?**



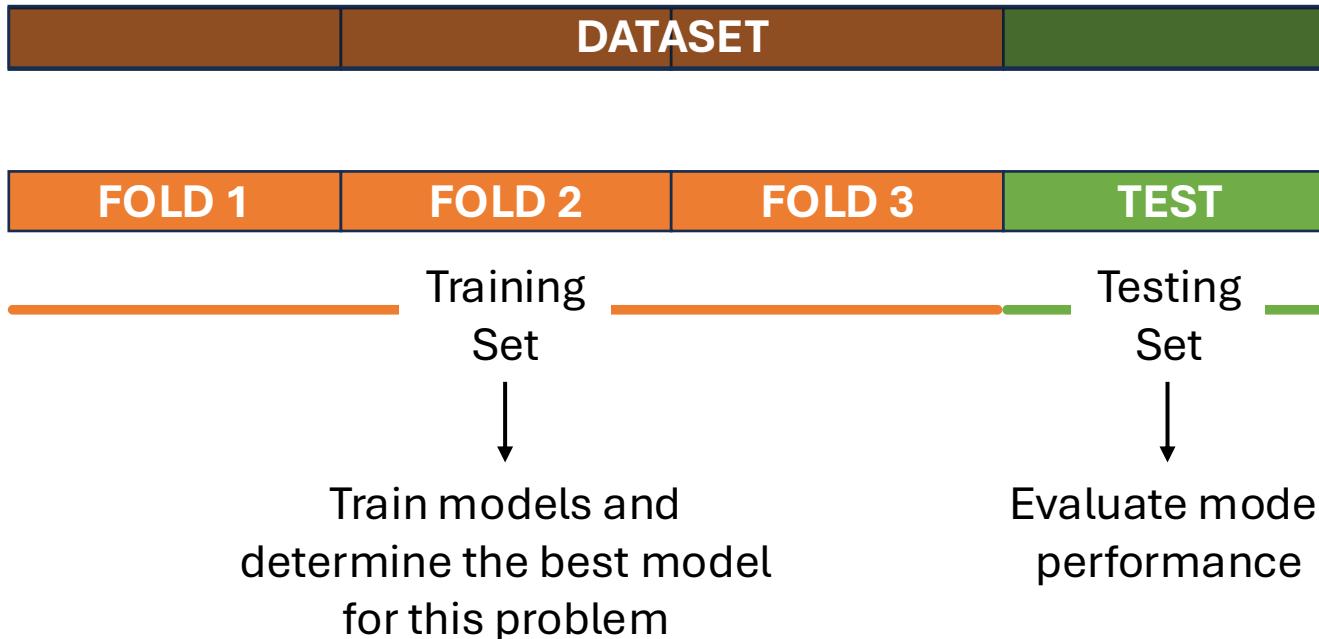
Temerty Centre for AI Research  
and Education in Medicine

UNIVERSITY OF TORONTO

# APPROACH TO EVALUATING CLINICAL AI STUDIES

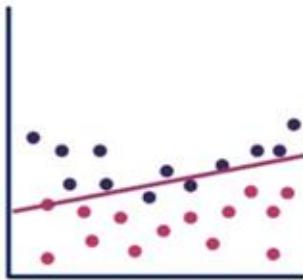


# TRAINING AND TESTING DATA



# MODEL FIT

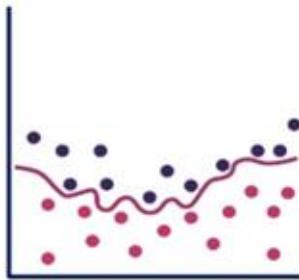
*Imagine the goal of this algorithm is to differentiate between purple & pink*



**Underfitting**



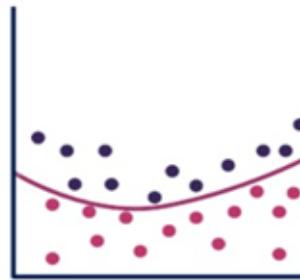
Poor performance  
on testing and  
training data



**Overfitting**



Performs very well on  
the data that trained it –  
will not perform as well  
on new data



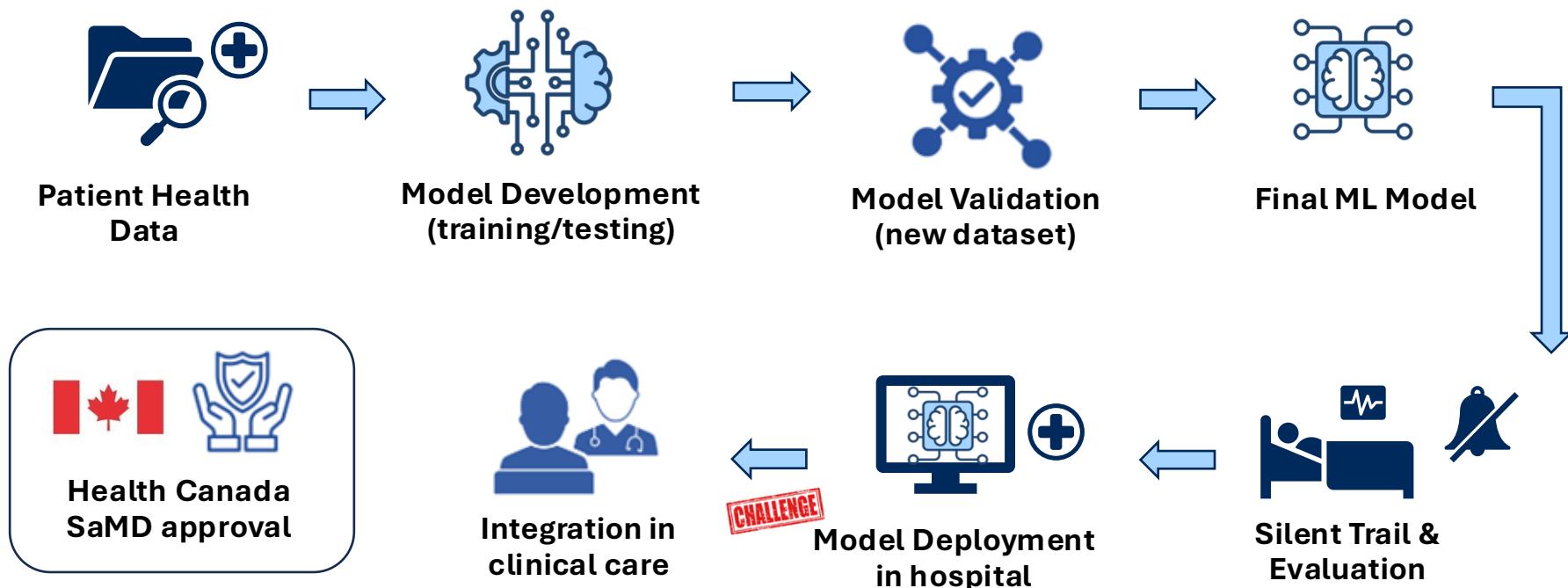
**Balanced**

Source: <https://www.r-bloggers.com/2022/09/how-to-avoid-overfitting/>

# **6. DEPLOYING AI TO CLINICAL SETTINGS**

How does AI become clinical? What do I need to consider in using a model clinically?

# HOW DOES ML BECOME CLINICAL?

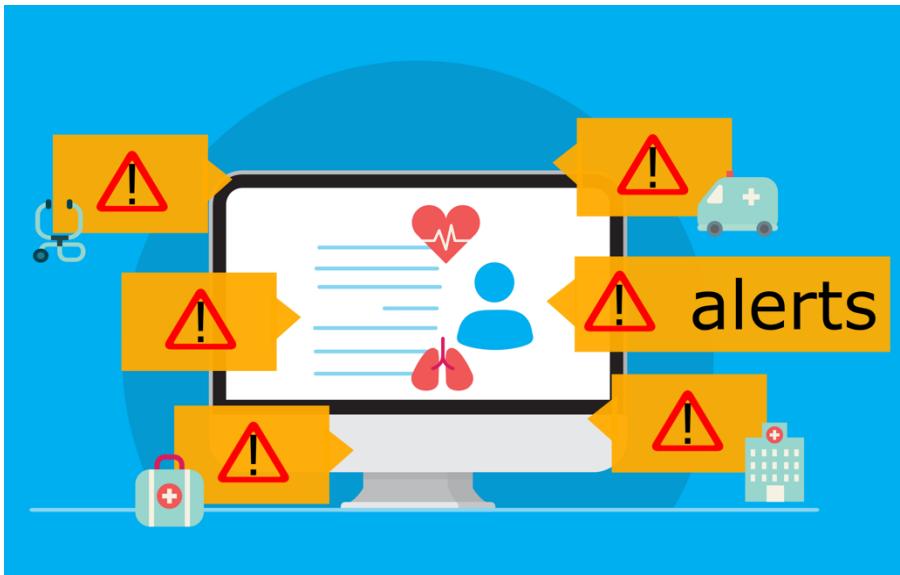


# DEPLOYMENT OF AI: BEYOND ALGORITHM

Theme	Key Considerations
Data	Are the data needed for the algorithm readily available and in an extractable format?
Infrastructure	Can relevant data be extracted in real time? Are there adequate infrastructure and computing resources available to host a cloud-based analytic and storage platform?
Interface	How will the clinical team be made aware of these predictions (e.g., dashboard)? How can the clinical team understand how the algorithm made this prediction?
End-Users	Which clinical team member(s) are most appropriate to receive the risk prediction? What is their level of trust in the AI algorithm? What is the risk of alert fatigue, burnout, or decrease in algorithmic compliance over time?
Clinical Context	Can a standardized action plan be implemented based on the risk prediction? Are there sufficient hospital resources to act on risk prediction?
Monitoring	How will clinical end-users or hospital leadership know whether the model continues to perform well? What happens if algorithm's performance degrades over time? What happens if compliance of clinical end-users degrades over time?

Source: Kwong et al. 2024 <https://www.nature.com/articles/s41746-024-01066-z/tables/1>

# CLINICAL EXAMPLE OF ML MODEL



**ML Model:** In clinic, supervised deep learning model implemented this year to predict risk of hospital 30-day re-admission for patients with heart failure.

**Patient:** 75 M (diabetic, high BMI, previous MI) presents to clinic for follow-up after 7-day admission for worsening heart failure (dyspnea, edema). Stabilized by changes in medications – changes were well tolerated.

**Challenge:** Patient appears well. Model predict high risk of re-admission. No imaging facilities in local community.

# WHAT IS T-CAIREM?

**T-CAIREM:** Temerty Center for Artificial Intelligence Research and Education in Medicine

# ORIGIN OF T-CAIREM

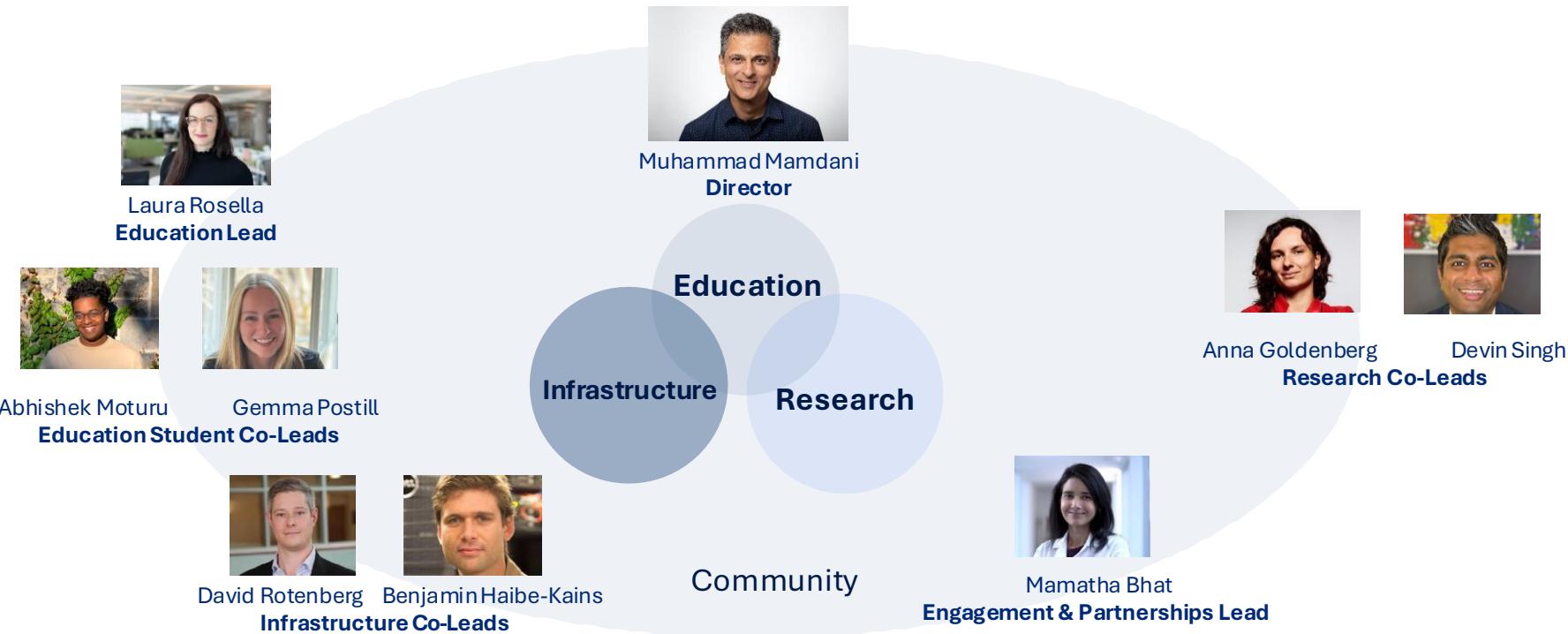
**The Temerty Centre for Artificial Intelligence Research and Education in Medicine (T-CAIREM)**

Launched in October 2020 thanks to a generous donation from the Temerty family.

Our mission is to transform health through AI.

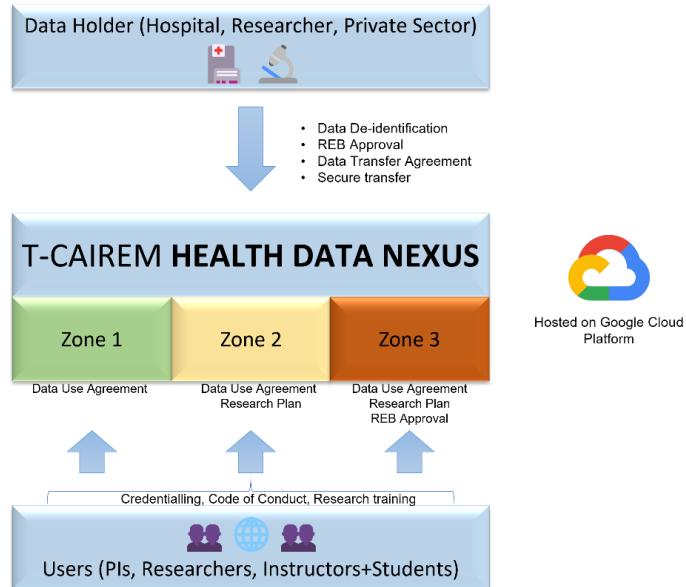


# T-CAIREM: STRUCTURE AND LEADERSHIP



# HEALTH DATA NEXUS

**One part data access platform, one part analytics environment**



# THE T-CAIREM COMMUNITY



- Over 1200 members**
- 800 faculty members**
- 460 students**
- 25 Canadian university partners**
- 2 international university partners**
- Over 90 Canadian hospital and research institutions**

# THE GLOBAL AI IN MEDICINE COMMUNITY



Winner of the 2022 AIMed  
Hospital/Institution of the Year Award

T-CAIREM is a member of the Alliance of Centres of Artificial Intelligence in Medicine (ACAIM) along with more than 50 of the world's leading AI in medicine centres.

