Introduction to Machine Learning

Introduction to ML Concepts, Regression, and Classification

Instructor: Pat Virtue

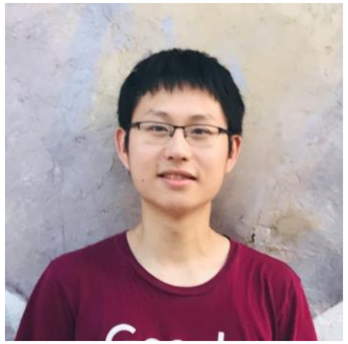
Course Staff

Teaching Assistants

Alex
Singh

Annie
Hu

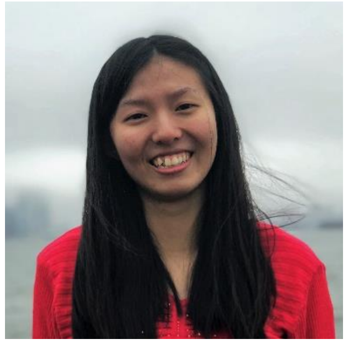
George
Brown



Haoran
Fei

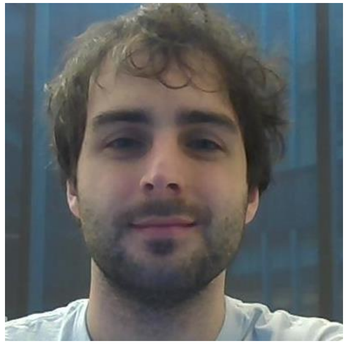
Michell
Ma

Nidhi
Jain



Vicky
Zeng

Education Associate



Daniel
Bird

Instructor

Pat
Virtue

Course Information

Website: https://www.cs.cmu.edu/~10315

Canvas: canvas.cmu.edu

Gradescope: gradescope.com

Communication: piazza.com

E-mail: dpbird@andrew.cmu.edu

pvirtue@andrew.cmu.edu

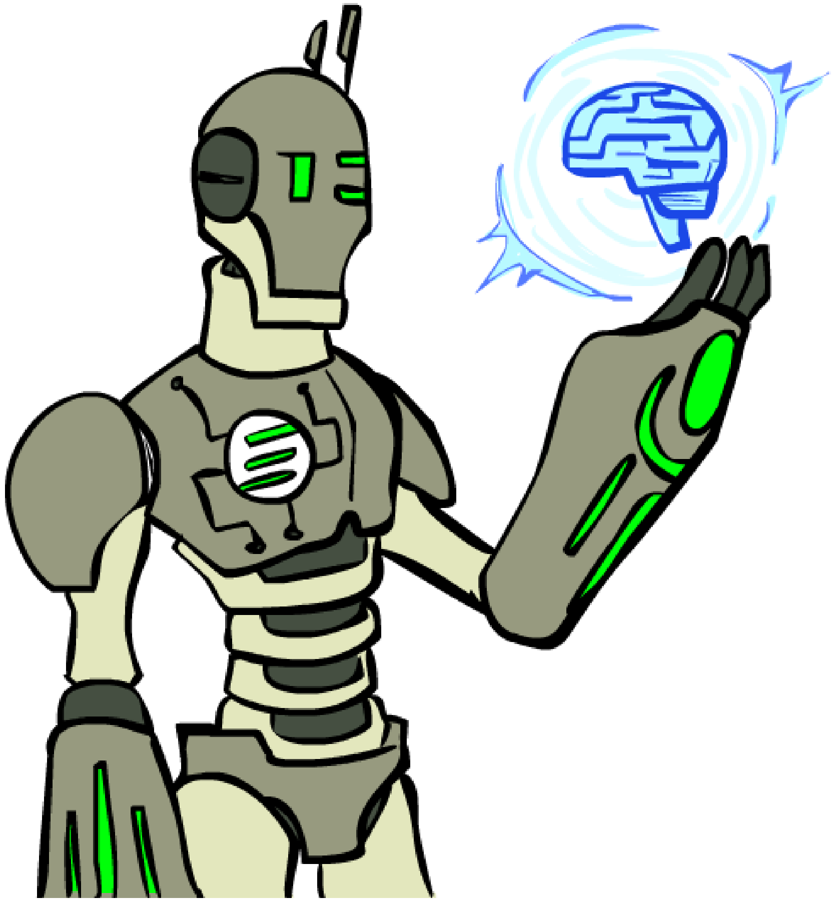
Today

What is AI/ML?

A brief history of AI/ML

Some logistics

Introduction to important ML concepts that we'll use throughout the semester



Images: ai.berkeley.edu

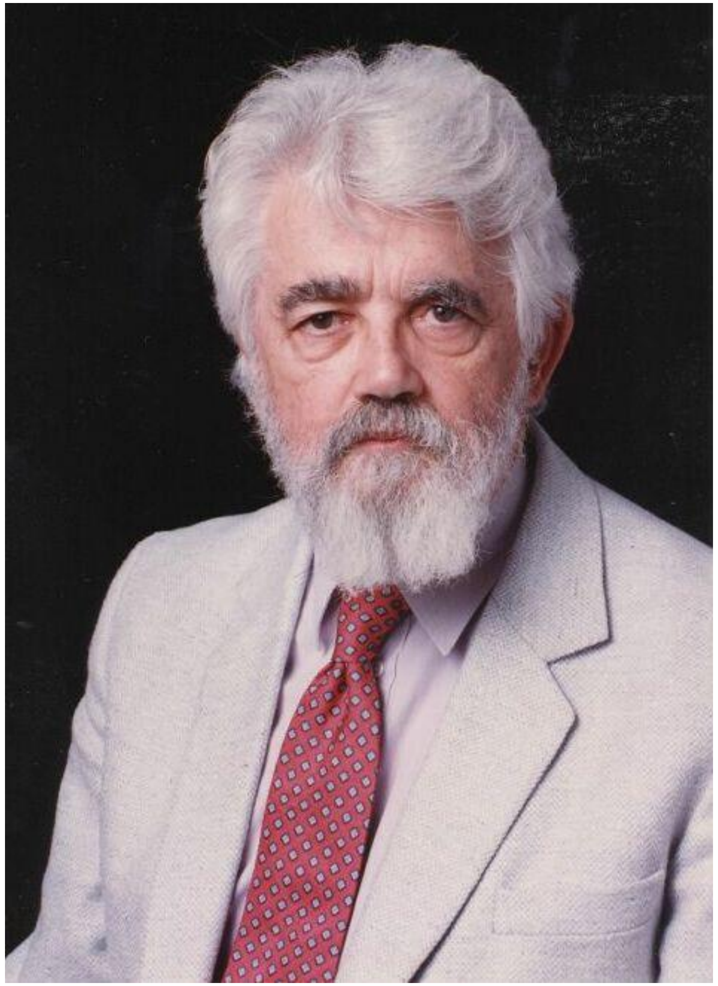
AI Definition by John McCarthy

What is artificial intelligence

- It is the science and engineering of making intelligent machines, especially intelligent computer programs

What is intelligence

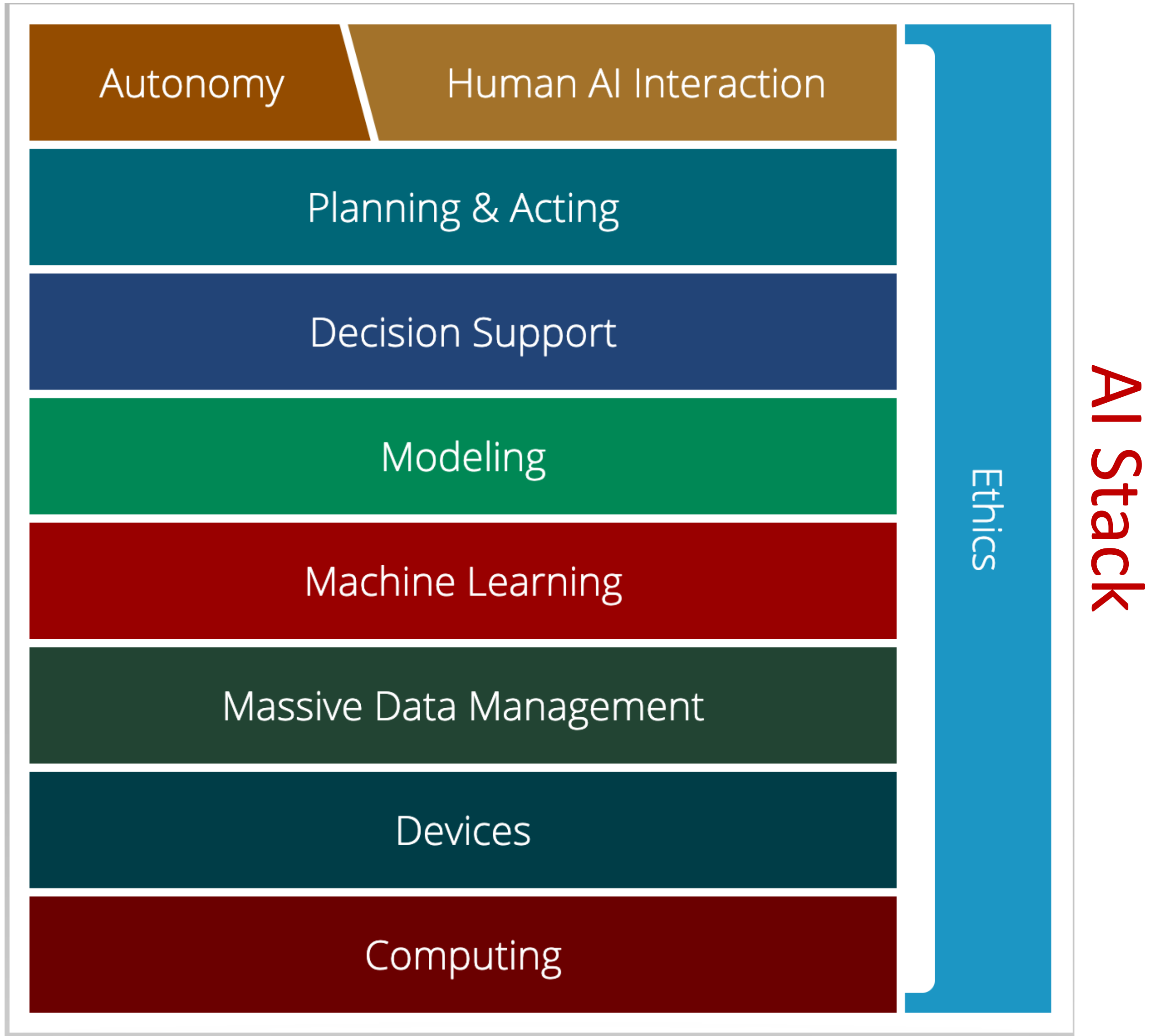
- Intelligence is the computational part of the ability to achieve goals in the world



http://www-formal.stanford.edu/jmc/whatisai/whatisai.html

AI Stack for CMU AI

"AI must understand the human needs and it must make smart design decisions based on that understanding"

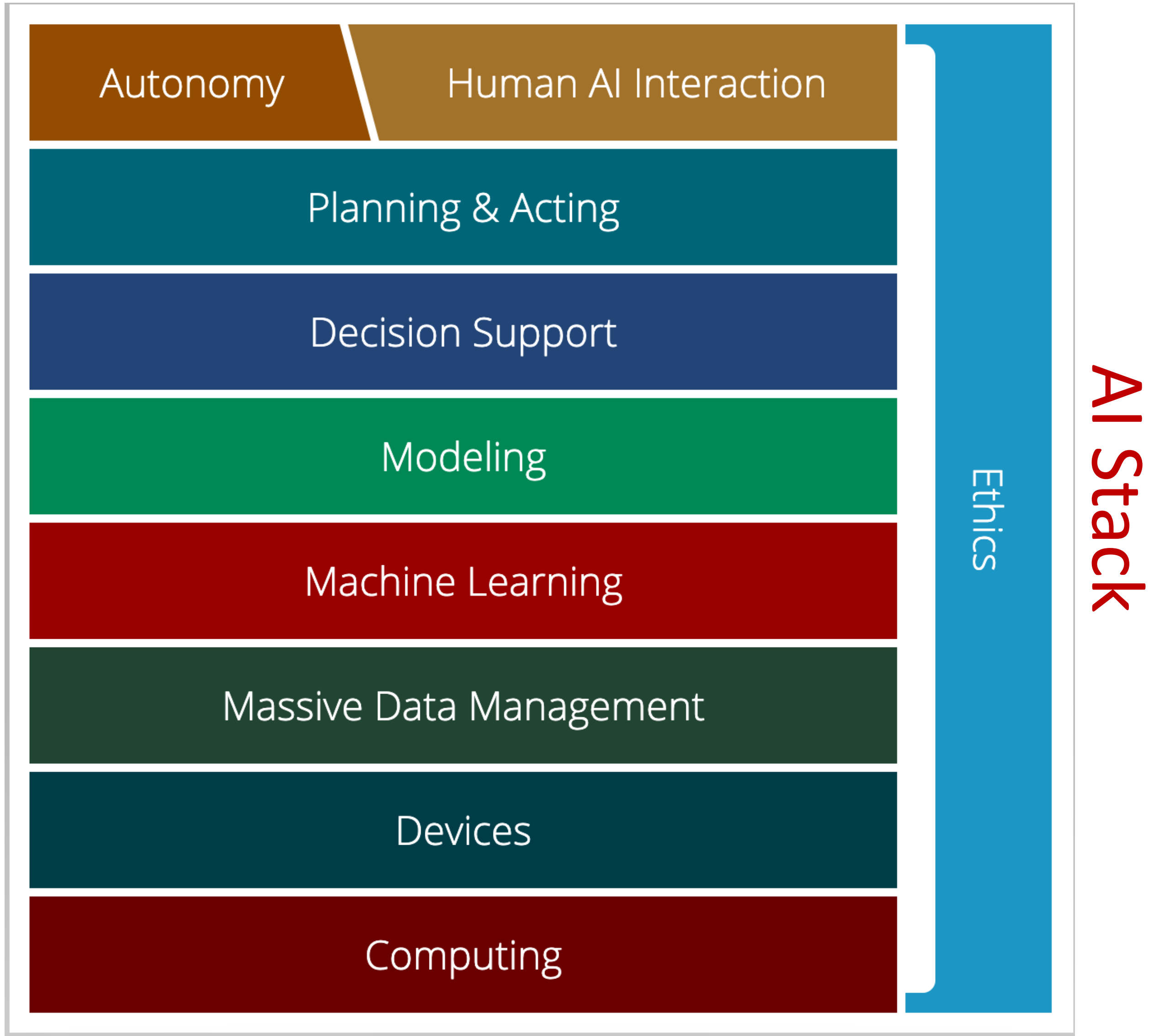


https://ai.cs.cmu.edu/about

AI Stack for CMU AI

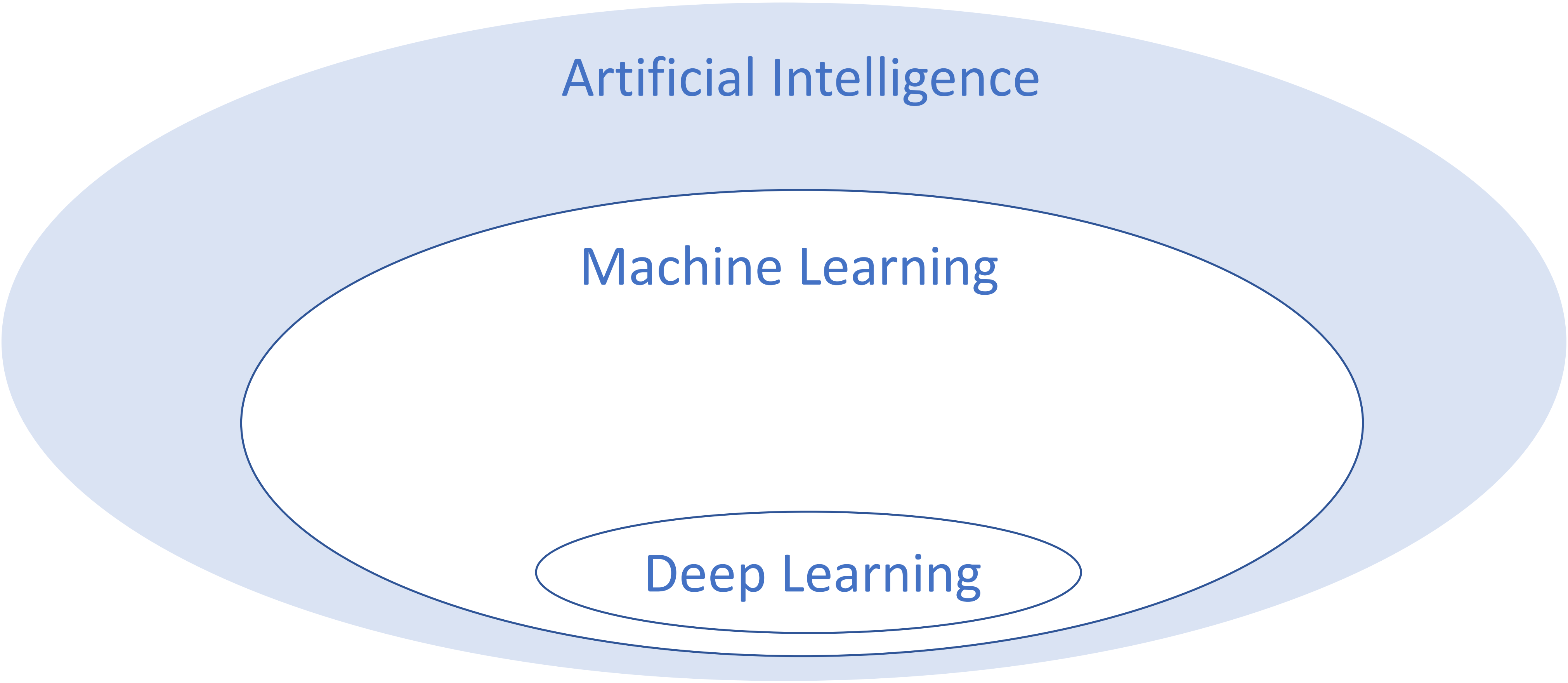
“Machine learning focuses on creating programs that learn from experience.”

“It advances computing through exposure to new scenarios, testing and adaptation, while using pattern- and trend-detection to help the computer make better decisions in similar, subsequent situations.”

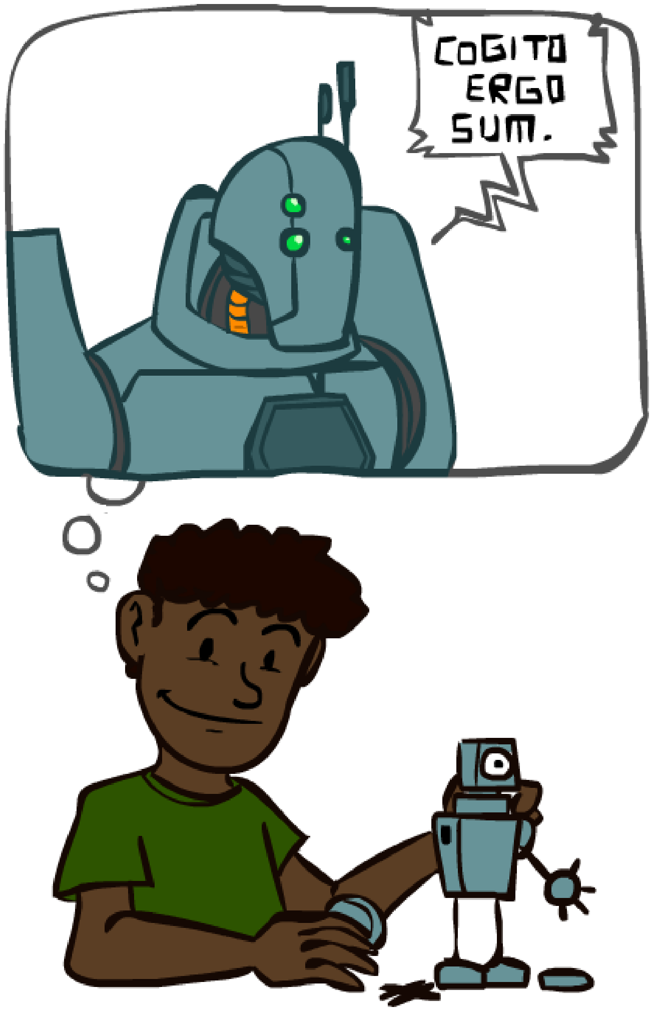


https://ai.cs.cmu.edu/about

Artificial Intelligence vs Machine Learning?

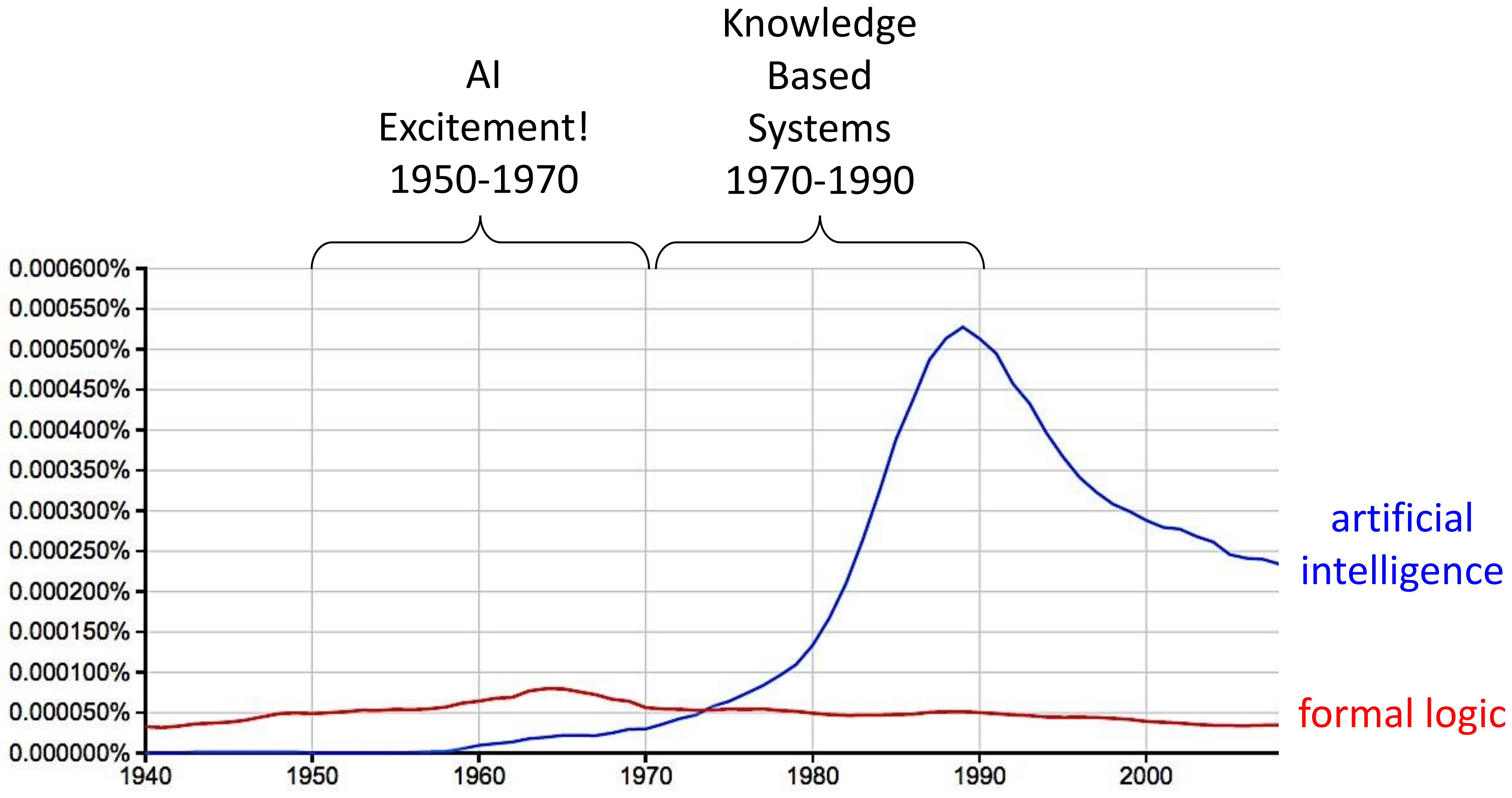


A Brief History of AI



Images: ai.berkeley.edu

A Brief History of AI



https://books.google.com/ngrams

What went wrong?

Dog

- Barks
- Has Fur
- Has four legs

Buster

✓

✓

✗

Knowledge-based Systems

Write programs that simulate how people do it.

Problems:

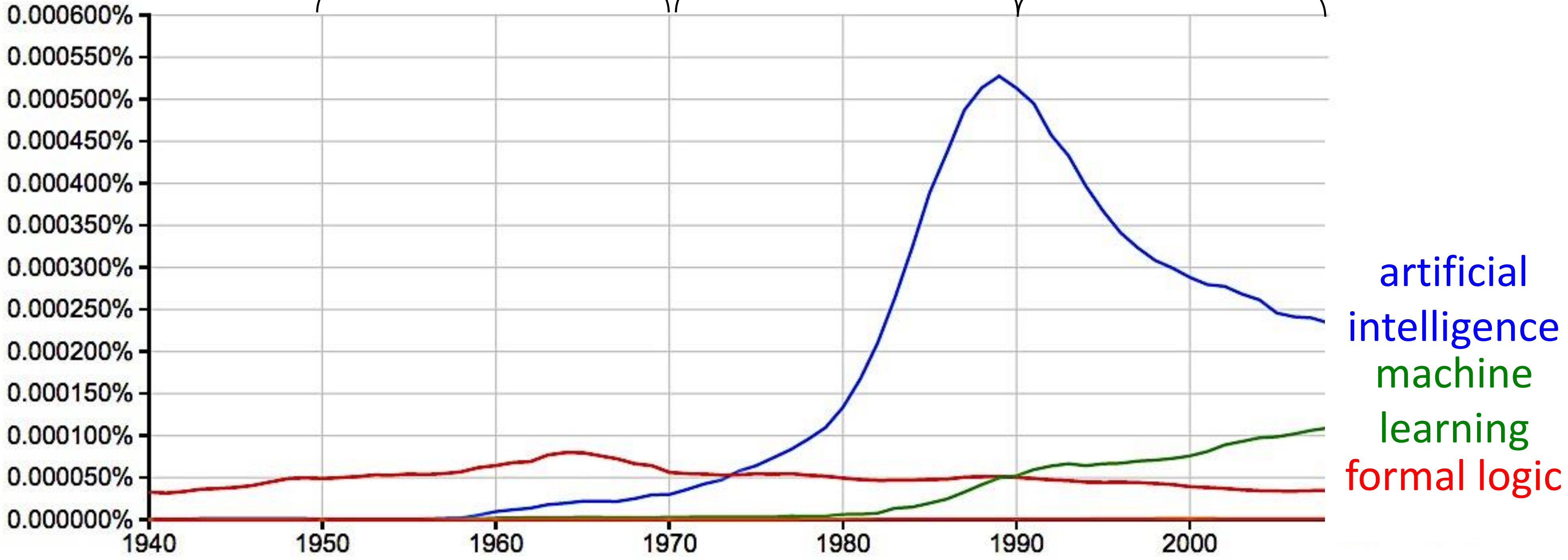
- Will never get better than a person
- Requires deep introspection
- Sometimes requires experts ("expert systems", "knowledge elicitation")
- Often, we don't know how we do things (e.g. ride bicycle)
  - Difference between knowing and knowing-how-we-know
- Sometimes we *think* we know, but we're wrong

A Brief History of AI

AI Excitement! 1950-1970

Knowledge Based Systems 1970-1990

Statistical Approaches 1990--



https://books.google.com/ngrams

Statistical Methods

Write programs that learn the task from examples

+ You don’t need to know how to do it yourself

+ Performance (should) improve with more examples

But:

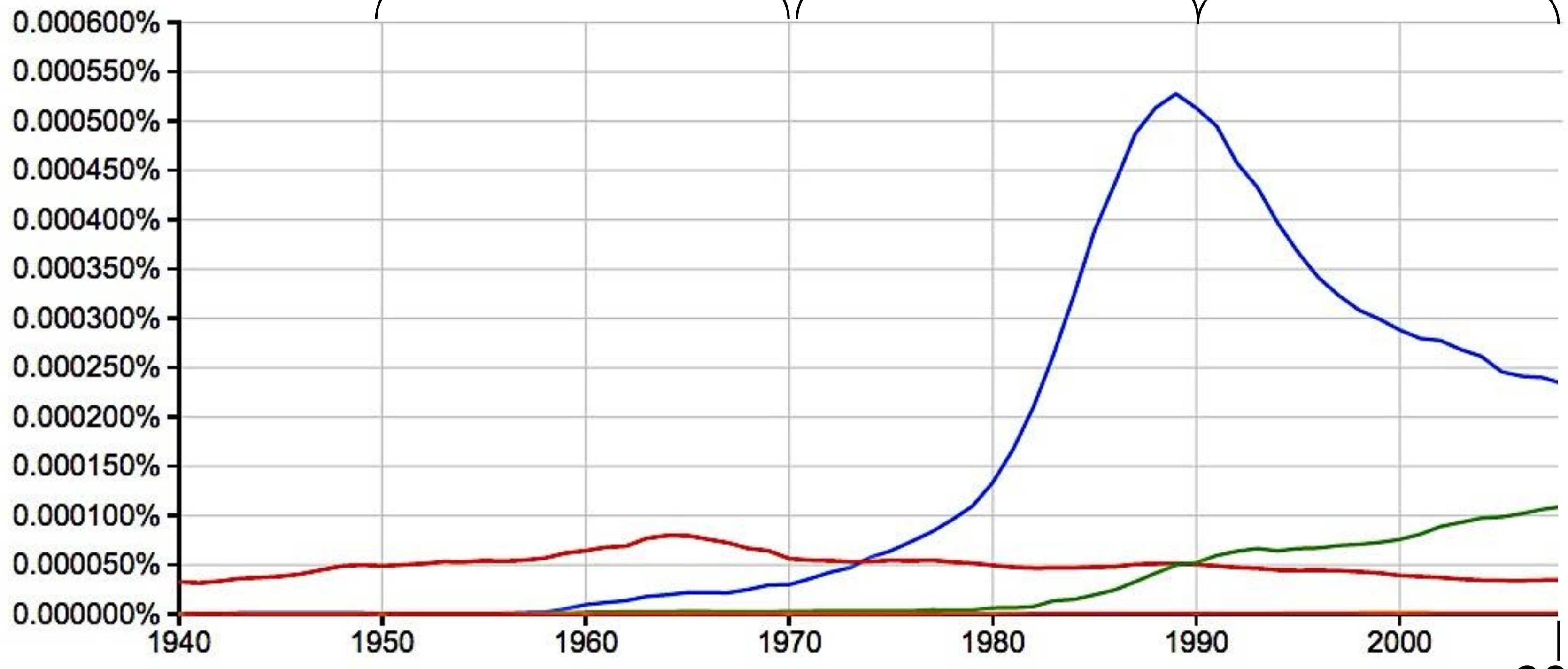
- Need lots of examples!
- When it finally works, you may not understand how

A Brief History of AI

AI Excitement! 1950-1970

Knowledge Based Systems 1970-1990

Statistical Approaches 1990--



artificial intelligence

machine learning

formal logic

2008

https://books.google.com/ngrams

A Brief History of AI

1940-1950: Early days

- 1943: McCulloch & Pitts: Boolean circuit model of brain
- 1950: Turing's "Computing Machinery and Intelligence"

1950—70: Excitement: Look, Ma, no hands!

- 1950s: Early AI programs, including Samuel's checkers program, Newell & Simon's Logic Theorist, Gelernter's Geometry Engine
- 1956: Dartmouth meeting: "Artificial Intelligence" adopted

1970—90: Knowledge-based approaches

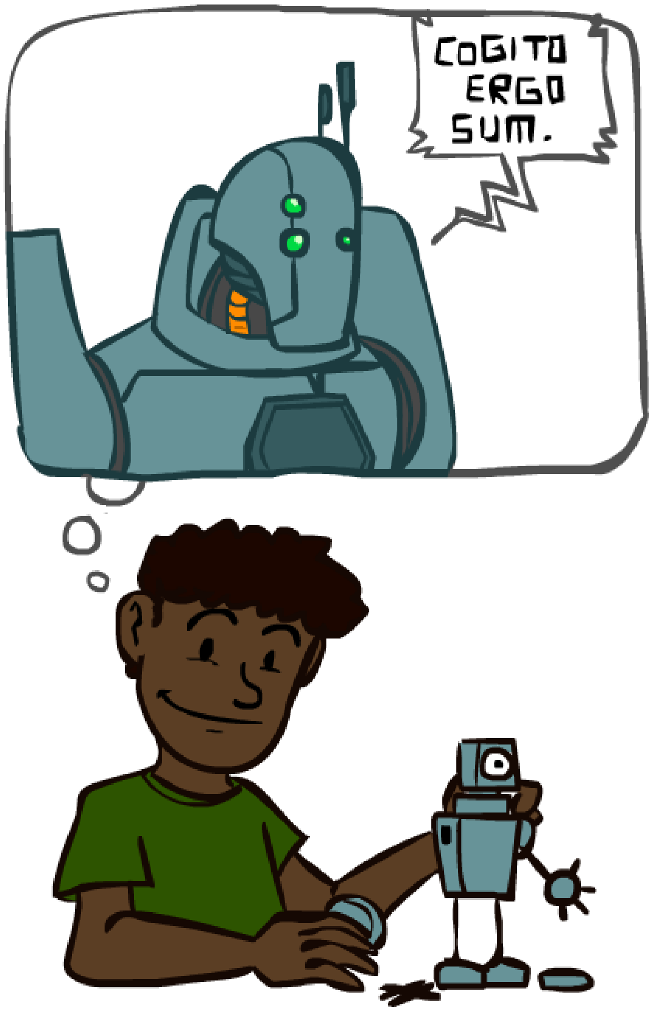
- 1969—79: Early development of knowledge-based systems
- 1980—88: Expert systems industry booms
- 1988—93: Expert systems industry busts: "AI Winter"

1990—: Statistical approaches

- Resurgence of probability, focus on uncertainty
- General increase in technical depth
- Agents and learning systems… "AI Spring"?

2012—: Deep learning

- 2012: ImageNet & AlexNet



Images: ai.berkeley.edu

Machine Learning and Statistics

Statistics is also about learning from data

Statistics has been around from much longer!

What's the difference?

Until the mid 1990s:

Statistics:

- A branch of mathematics
- Emphasized rigor, correctness, provable properties ("is it correct?")
- Was not very concerned with scaling
  - Not much awareness of computational complexity

Machine Learning:

- A branch of Computer Science / AI
- Focus on heuristics, making things work in practice ("does it work?")
- Not much awareness of statistical theory

Slide: CMU ML, Roni Rosenfeld

Machine Learning and Statistics

From the mid 1990s:

The two fields have effectively merged

- Carnegie Mellon has led the way!

ML is now often called “Statistical Machine Learning”

- There is very little non-statistical ML today

The Machine Learning Framework

Formalize the task as a mapping from input to output

- Task examples will usually be pairs: (input, correct_output)

Formalize performance as an error measure

- or more generally, as an objective function (aka Loss function)

Examples:

- Medical Diagnosis
  - mapping input to one of several classes/categories (aka classification)
- Predict tomorrow's Temperature
  - mapping input to a number (aka regression)
- Chance of Survival: From patient data to p(survive >= 5 years)
  - mapping input to probability (aka logistic regression)
- Driving recommendation
  - mapping input into a plan (aka Planning)

Slide: CMU ML, Roni Rosenfeld

Choices in ML Problem Formulation

Often, the same task can be formulated in more than one way:

Ex. 1: Loan applications

- creditworthiness/score (regression)
- probability of default (logistic regression)
- loan decision (classification)

Ex. 2: Chess

- Nature of available training examples/experience:
  - expert advice (painful to experts)
  - games against experts (less painful but limited, and not much control)
  - experts' games (almost unlimited, but only "found data" – no control)
  - games against self (unlimited, flexible, but can you learn this way?)
- Choice of target function: board→move vs. board→score

Slide: CMU ML, Roni Rosenfeld

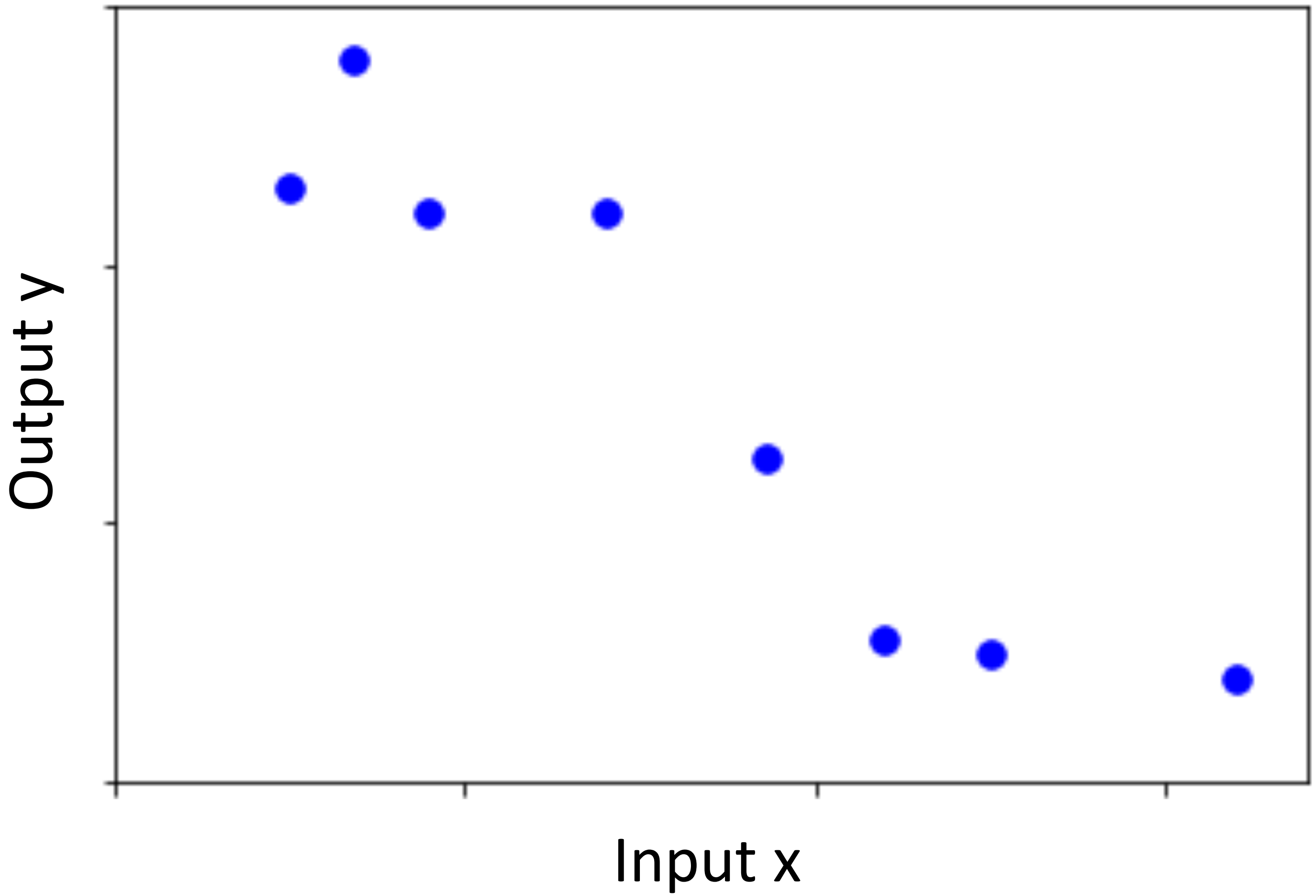
Machine Learning

We **cannot** learn from data

We **can** learn from data + *assumptions*

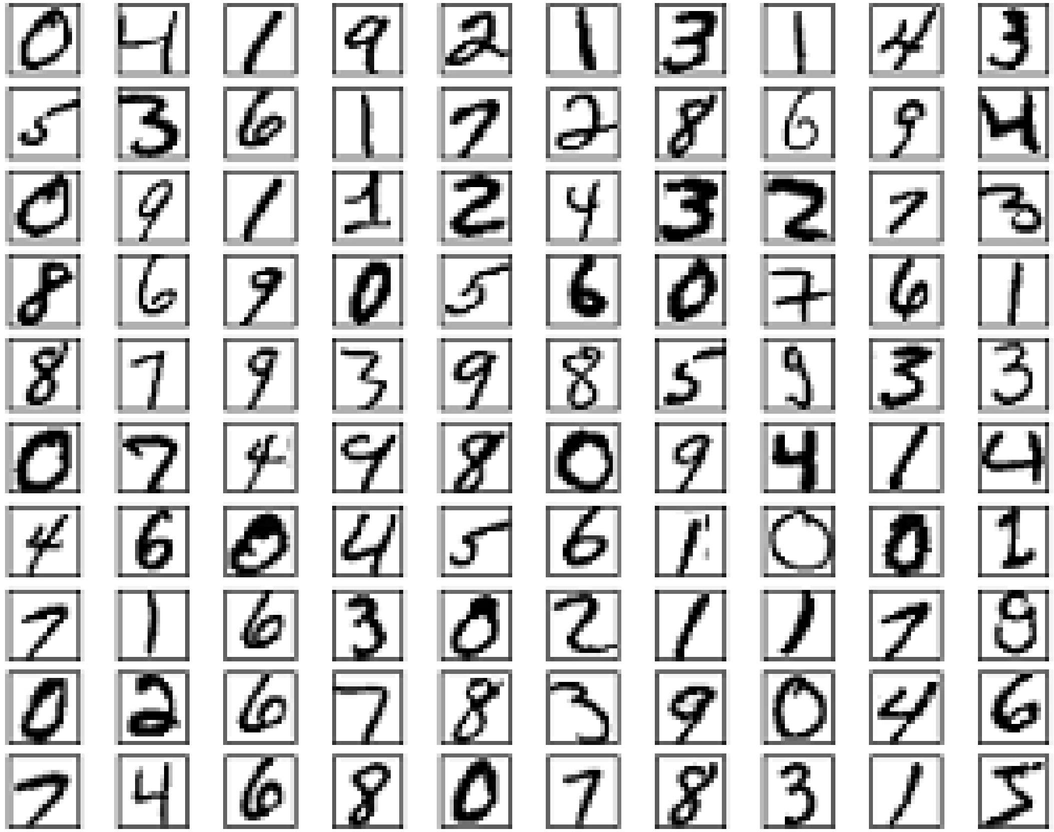
Assumptions

What assumptions do we make with this data?



Assumptions

What assumptions do we make with this data?



More robust examples: http://yann.lecun.com/exdb/lenet/index.html

Assumptions

Face dataset

Course Information

Website: https://www.cs.cmu.edu/~10315

Canvas: canvas.cmu.edu

Gradescope: gradescope.com

Communication: piazza.com

E-mail: dpbird@andrew.cmu.edu

pvirtue@andrew.cmu.edu

Course Information

- Course Levels and Course Scope
- Participation
- Video
- Waitlist
- Prerequisites
- Teamwork
- Mental health

Announcements

Recitation starting this Friday

- Recommended. Materials are fair game for exams
- One Section

Assignments:

- HW1 (online)
  - Released this week
  - Due next week

How to Approach a Machine Learning Problem

1. Consider your goal → definition of task **T**

- E.g. make good loan decisions, win chess competitions, …

2. Consider the nature of available (or potential) experience **E**

  - How much data can you get? What would it cost (in money, time or effort)?

3. Choose type of output **Y** to learn

- (Numerical? Category? Probability? Plan?)

4. Choose the Performance measure **P** (error/loss function)

5. Choose a representation for the input X

6. Choose a set of possible solutions **H** (hypothesis space)

- set of functions h: X ➔ Y
- (often, by choosing a representation for them)

7. Choose or design a learning algorithm

- for using examples (**E**) to converge on a member of **H** that optimizes **P**

Slide: CMU ML, Roni Rosenfeld, Tom Mitchel

Notation

Vocab: General ML Concepts

Data / examples / experience
- Input
- Output (labels)

Model

Parameters
- Model complexity

Hypothesis function
- Prediction

Error/loss, accuracy

Objective function

Global/local min/max

Training, validation, test set

Over (under) fitting

Classification

Regression

Supervised (unsupervised) learning

Vocab: Specific ML Concepts / Techniques

Linear model

Mean squared error

Gradient Descent

Stochastic Gradient Descent

Learning rate

Batch

Sigmoid

ReLU

Softmax

Cross entropy

Neural network