Math Foundations for ML

10-606

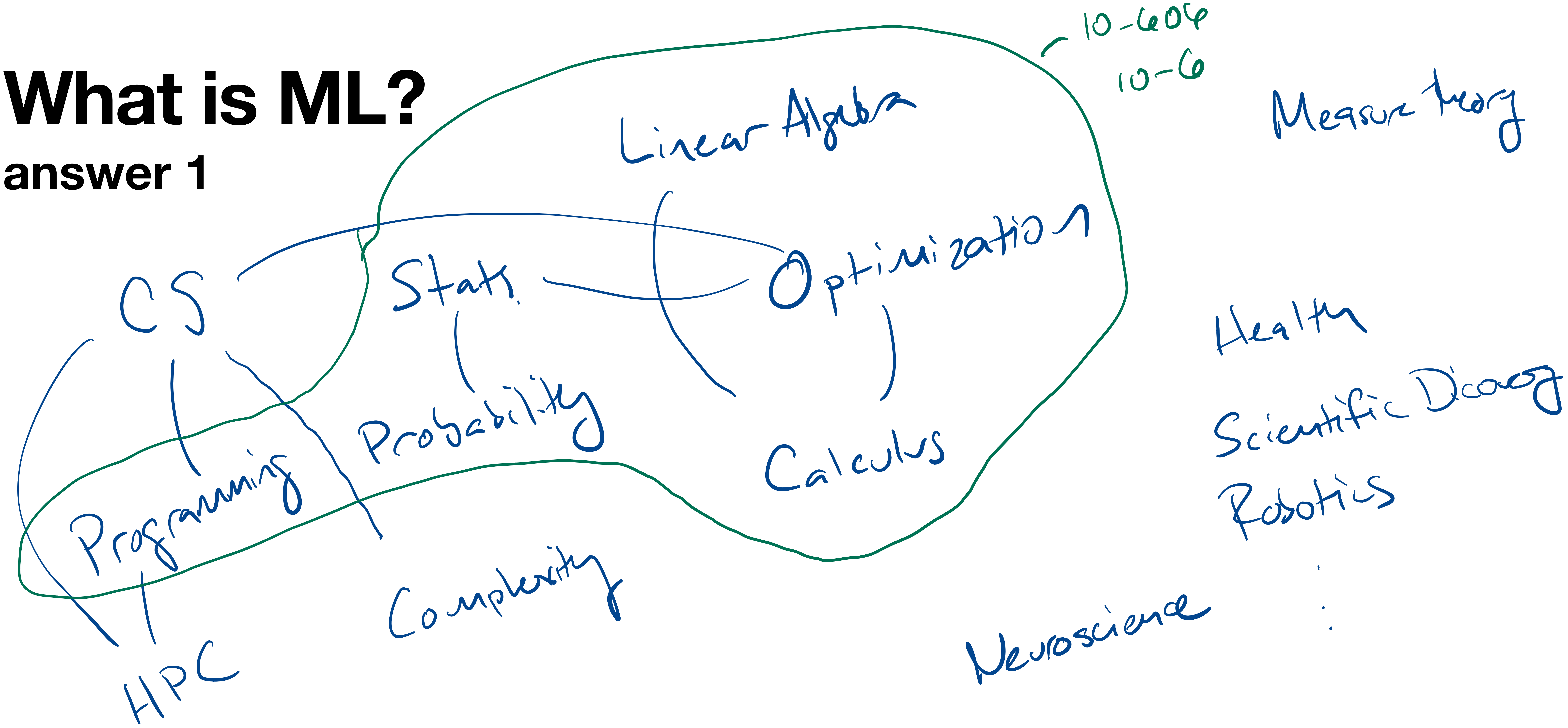
Geoff Gordon

About us

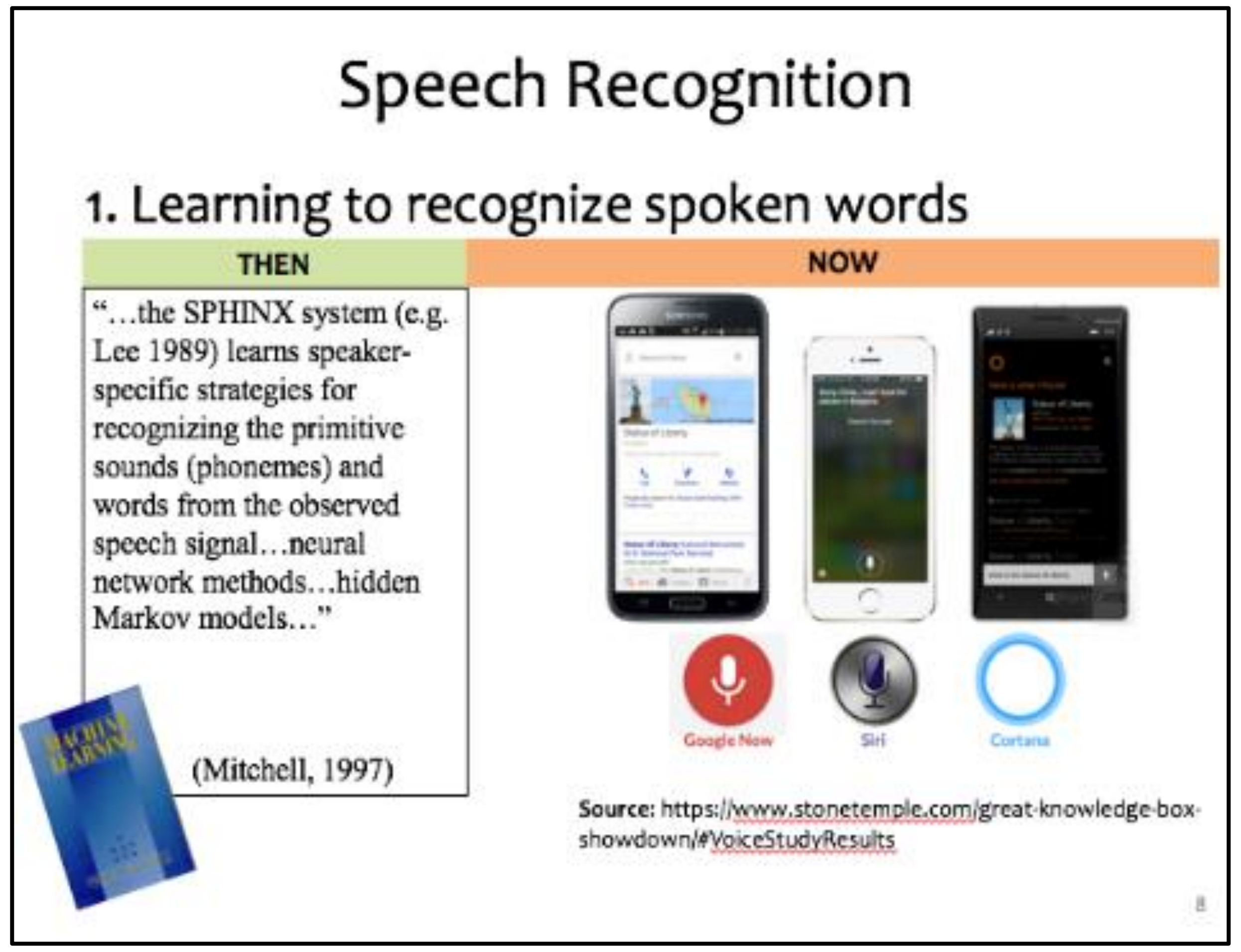
- Me: Geoff Gordon
- TAs:
  - Aditya Paul
  - Xiaoyu Xu
  - Aishwarya Jadhav

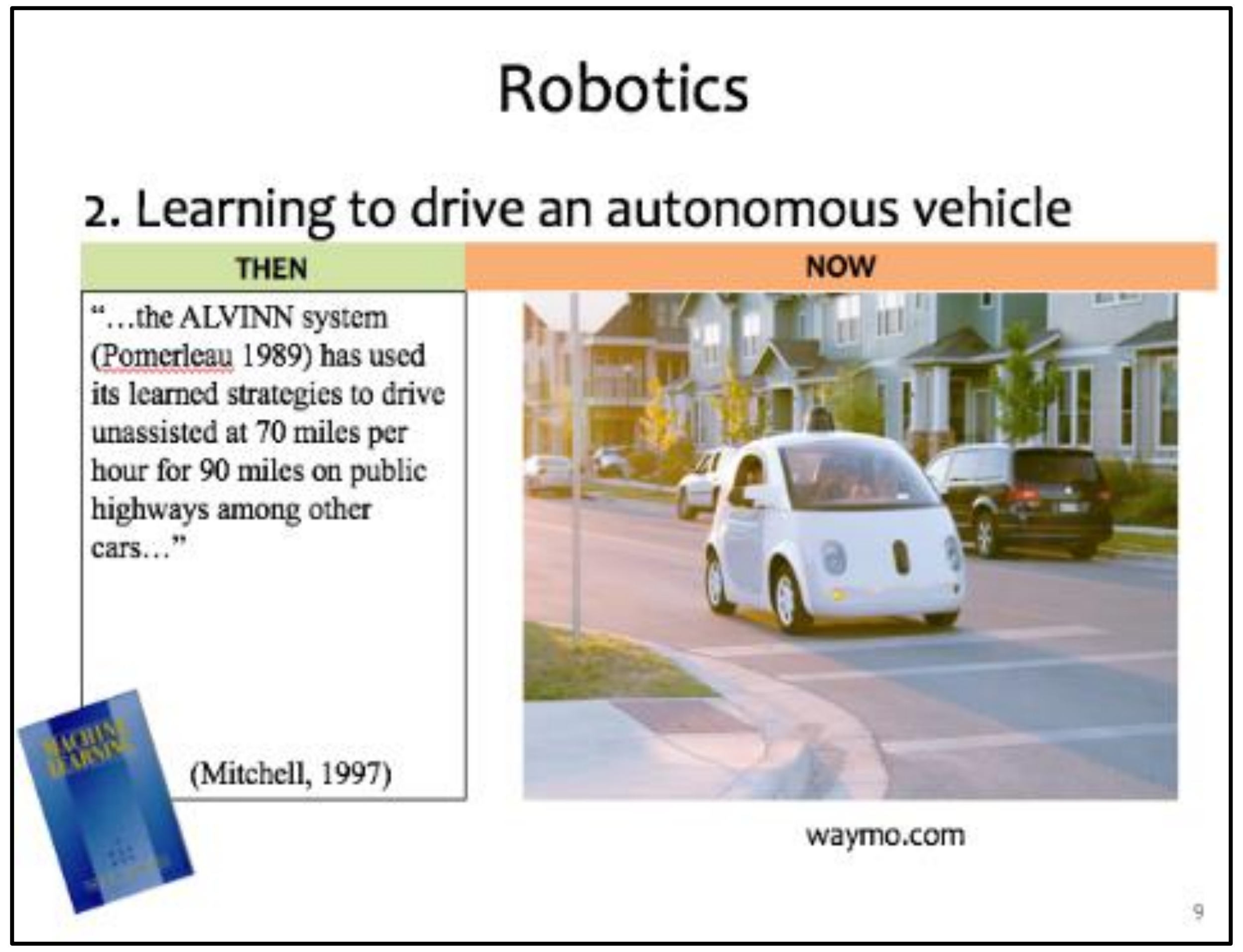
What is ML?

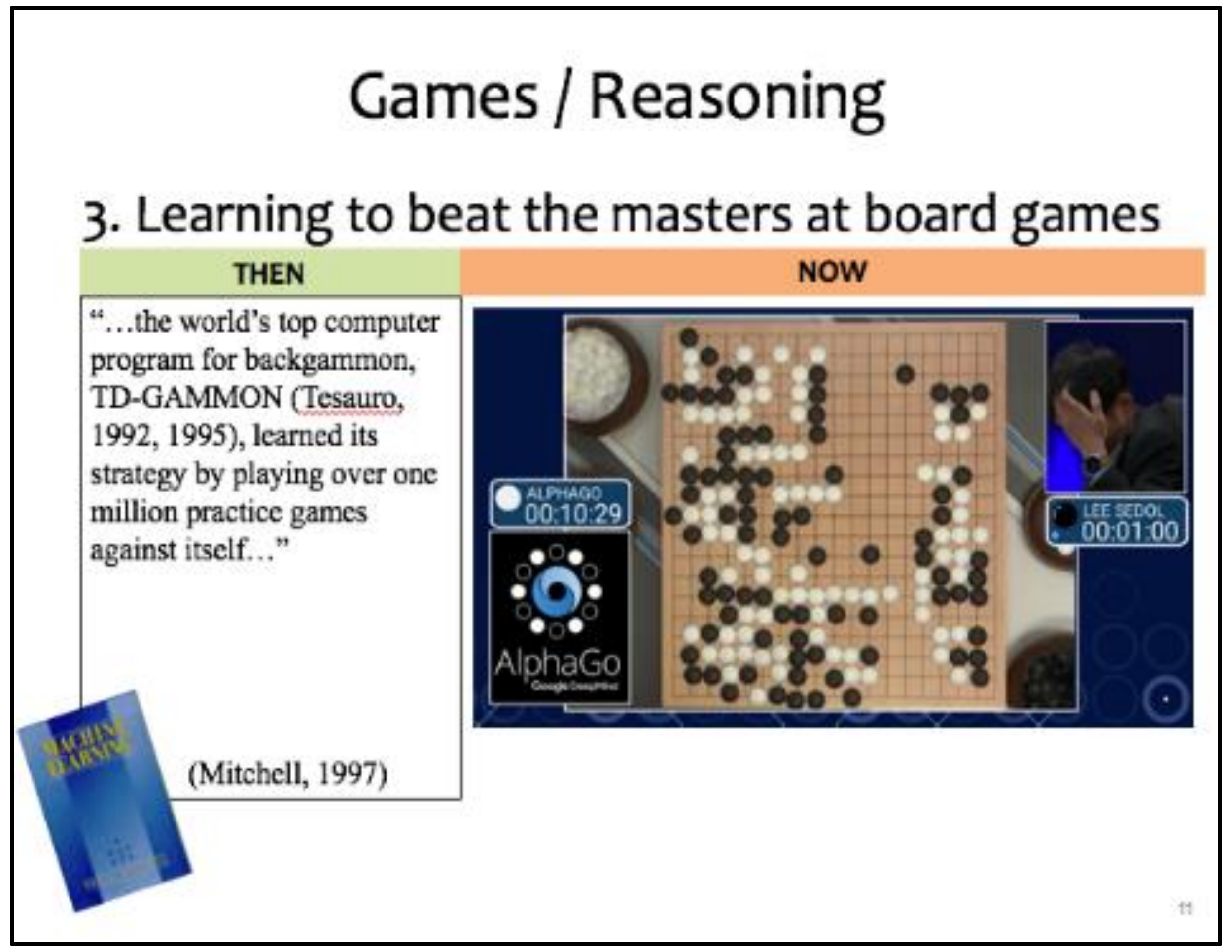
answer 1

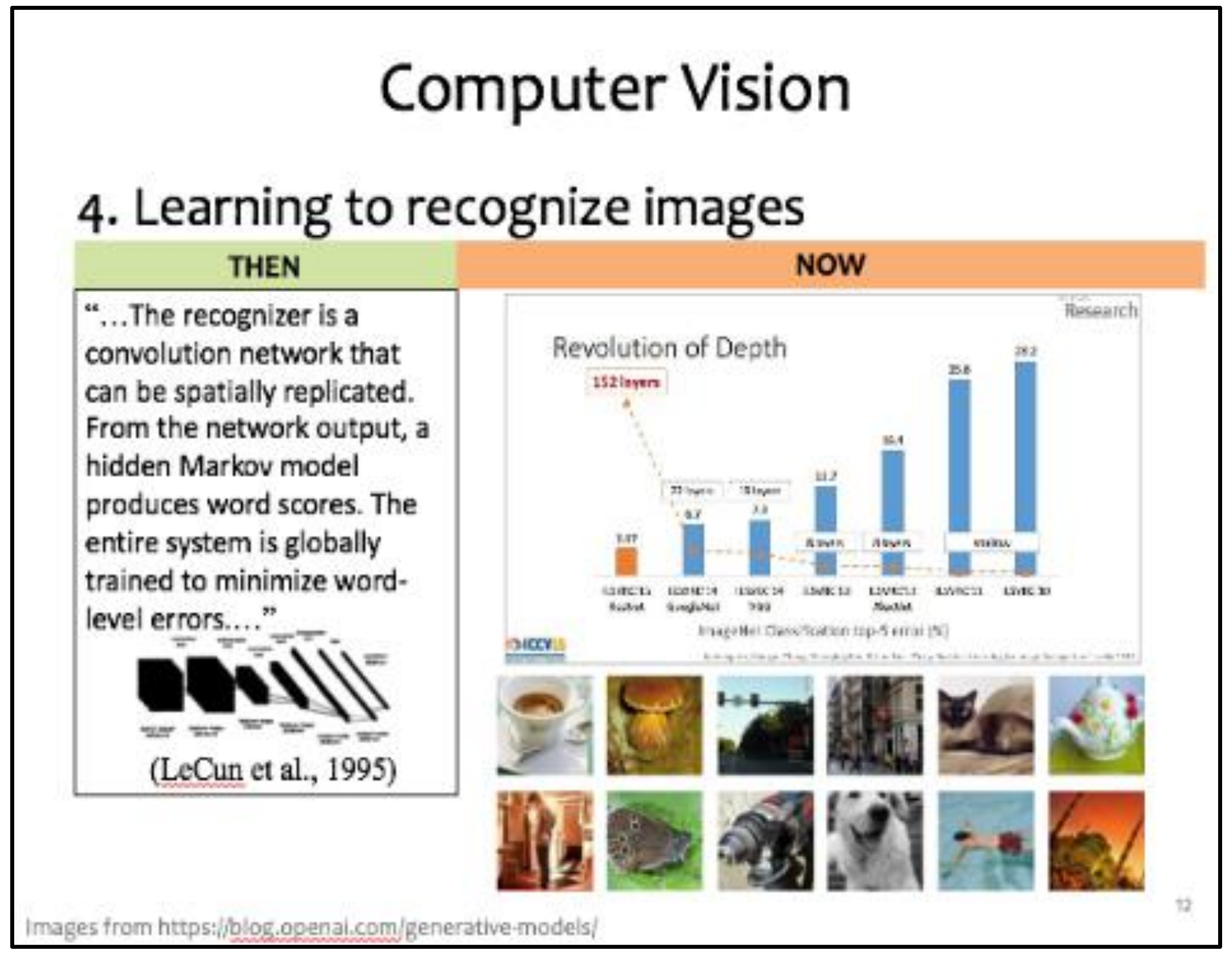


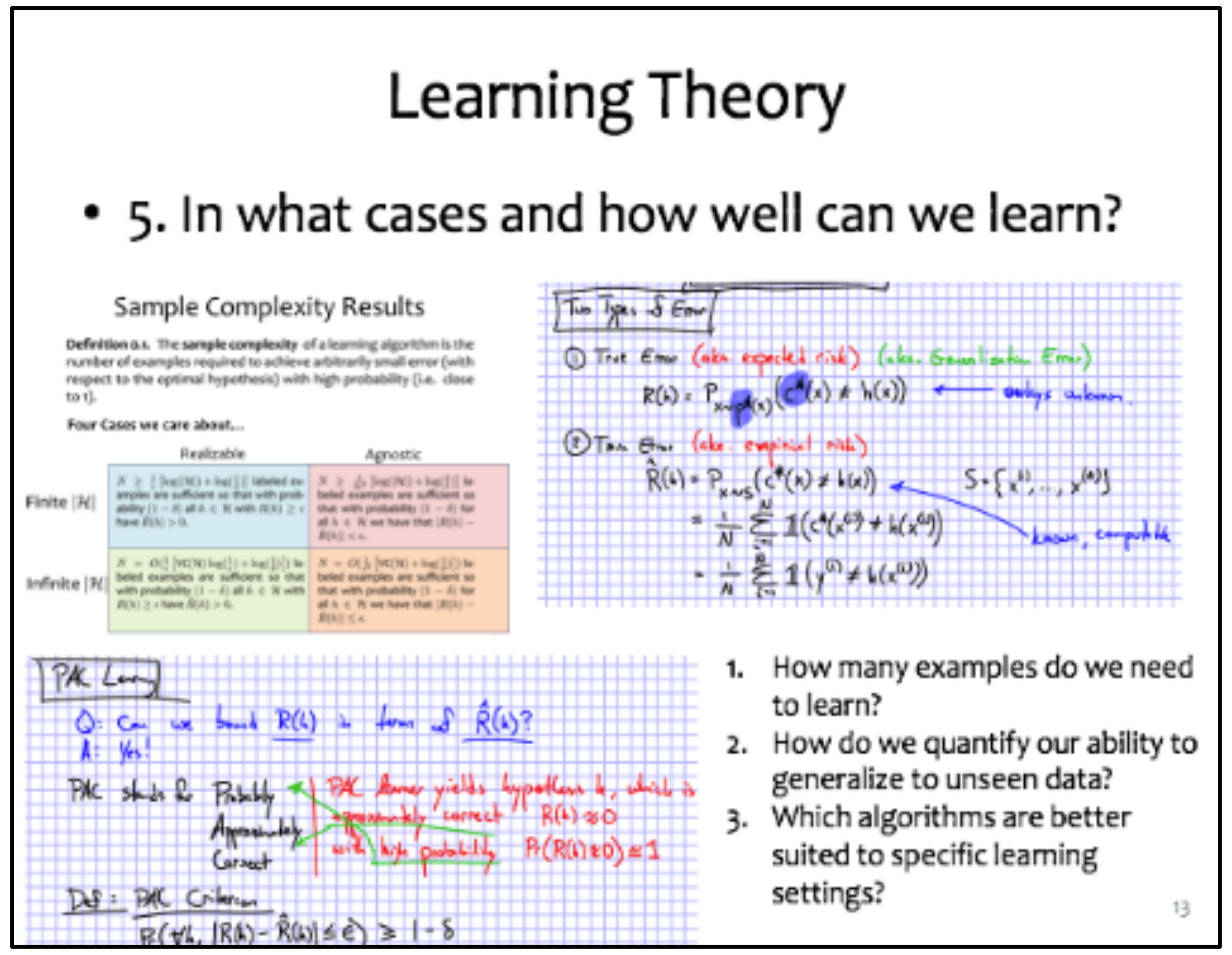
What is ML?
answer 2











credit: Matt Gormley

Why this course?

- Dual problem (derivation):

$$L(\mathbf{w}, b, \alpha) = \frac{1}{2}\mathbf{w}.\mathbf{w} - \sum_{j=1}^{n} \alpha_j \left[\left(\mathbf{w}.\mathbf{x}_j + b\right) y_j - 1\right]$$

$$\alpha_j \geq 0, \ \forall j$$

α - weights on training pts (n-dim problem)

Why this course?

Dual SVM – linearly separable case

- Dual problem (derivation):

$\max_{\alpha_j \geq 0} d(\alpha)$

$$\max_{\alpha} \min_{\mathbf{w},b} L(\mathbf{w}, b, \alpha) = \frac{1}{2}\mathbf{w}.\mathbf{w} - \sum_j \alpha_j \left[\left(\mathbf{w}.\mathbf{x}_j + b\right) y_j - 1\right]$$

$$\alpha_j \geq 0, \ \forall j$$

$$\frac{\partial L}{\partial w} = w - \sum_j \alpha_j x_j y_j$$

$$\rightarrow \frac{\partial L}{\partial \mathbf{w}} = 0 \qquad \Rightarrow \mathbf{w} = \sum_j \alpha_j y_j \mathbf{x}_j$$

$$\rightarrow \frac{\partial L}{\partial b} = 0 \qquad \Rightarrow \sum_j \alpha_j y_j = 0$$

$$\frac{\partial L}{\partial b} = \sum_j \alpha_j y_j$$

Dual SVM – linearly separable case

- Dual problem:

$$\sum_j \alpha_j b y_j = b \sum_j \alpha_j y_j = 0$$

$$\max_{\alpha} \min_{\mathbf{w},b} L(\mathbf{w}, b, \alpha) = \frac{1}{2}\mathbf{w}.\mathbf{w} - \sum_j \alpha_j \left[\left(\mathbf{w}.\mathbf{x}_j + b\right) y_j - 1\right]$$

$$\alpha_j \geq 0, \ \forall j$$

$$\Rightarrow \mathbf{w} = \sum_j \alpha_j y_j \mathbf{x}_j \qquad \Rightarrow \sum_j \alpha_j y_j = 0$$

$$\frac{1}{2} \sum_{ij} \alpha_i \alpha_j y_i y_j x_i \cdot x_j - \sum_j \alpha_j \left(\sum_i \alpha_i y_i x_i \cdot x_j \right) y_j + \sum_j \alpha_j$$

$$\sum_j \alpha_j - \frac{1}{2} \sum_{ij} \alpha_i \alpha_j y_i y_j x_i \cdot x_j =: d(\alpha)$$

credit: Aarti Singh (10-601)

Why this course?

4.2 Logistic Regression Implementation [30 points]

Implementation instructions.

1-Sentence Overview: You will be training a logistic regression model on the Homework 1 dataset by running gradient descent and then answering the written questions below.

Details: You will fill in the `logistic_regression.py` template and submit your complete file to Gradescope, where we will run your code against a suite of tests. Your grade will be automatically determined from the testing results. Since you get immediate feedback after …

credit: Aarti Singh, Geoff Gordon (10-601)

Course page

- https://www.cs.cmu.edu/~ggordon/10606s22/syllabus-and-lecture-outline.html

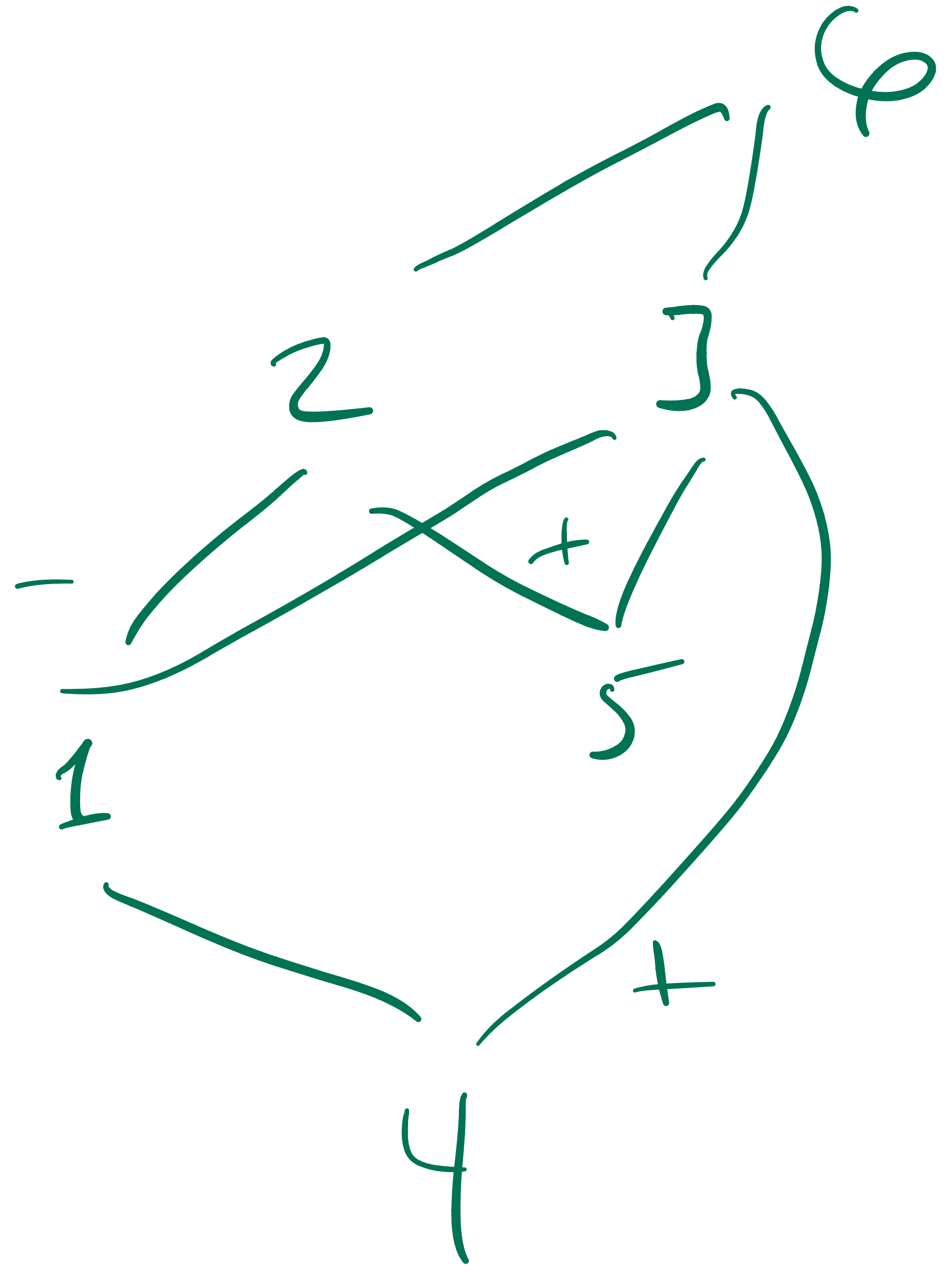
Help one another learn

- https://xkcd.com/1053/

Formal systems

0, 1, 2, 3, 4, 5, 6

+, −



Partial derivatives

Integrals

Car simulator

Gaussian elim

PM

propositional logic

classical

intuitionist ...

↳ ε

expressions: [abcde]*

rules:

erase ab

erase cd

erase e

a [

b]

c (

d)

e anything else

Set builder notation

$\{R, G, B\}$

$\{\} \rightarrow \emptyset$

$\{2x \mid x \in \mathbb{Z}\}$

↑ expr ↑ property

$\{2x \mid x \in \mathbb{Z},\ x > 0\}$

AND

$\{x \in \mathbb{Z} \mid x \% 2 = 0\}$

one simple property

{x**2 for x in range(4)}

↕

$\{x^2 \mid x \in \{0, 1, 2, 3\}\}$