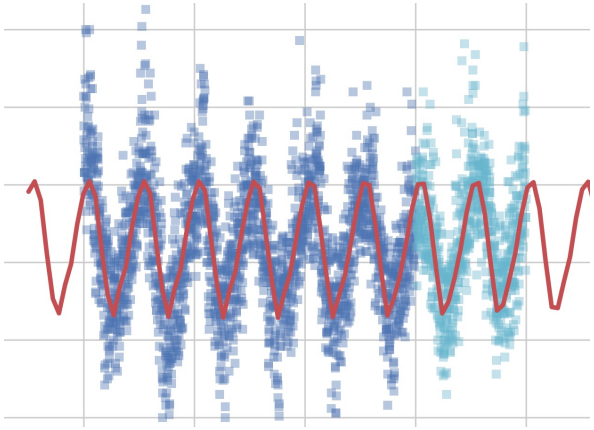
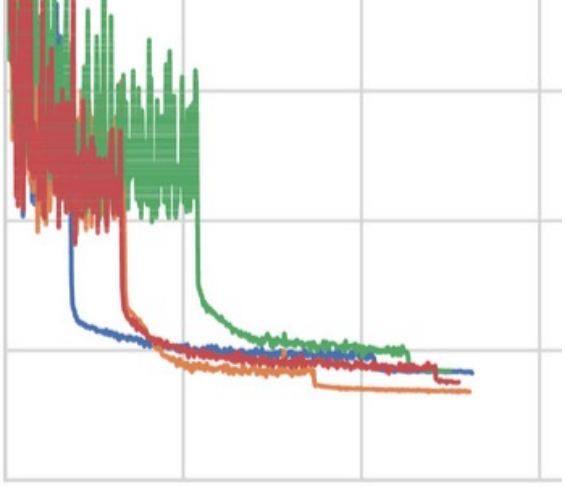




CS 135
Introduction to Machine Learning

Prof. Michael C. Hughes ("Mike")

Fall 2023, First day of class

As you join, please check out and have open all class:

* **Website:** https://www.cs.tufts.edu/cs/135/2023f/
  Read syllabus, skim schedule, waitlist info, etc.

Many slides attributable to: Emily Fox (UW), Finale Doshi-Velez (Harvard), Erik Sudderth (UCI), and Marty Allen & Liping Liu (Tufts)

Today's Agenda

- Who is teaching?
- Why take this course?
- What is Machine Learning?
- What skills/concepts will we learn?
- How will we spend our time?

Labs: Intro to NumPy

Who is teaching?

- Prof. Mike Hughes (www.michaelchughes.com)
  - Please call me “Mike”
    - (if you insist, I’ll answer to “Prof. Mike” or “Prof. Hughes”)

- TA Staff
  - Preetish Rath
  - Si Liu
  - Sipei Li
  - Elizabeth Cucuzzella

Quick Intros
- Say your name
- Degree program?
- Favorite thing about ML?

How to get help from staff?

- Come to office hours!

https://www.cs.tufts.edu/cs/135/2023f/office_hours.html

- Ask on Piazza (enrolled students only)

Q: Why should you take this course?

A: Machine Learning is everywhere! Those who know how to wield it effectively and *responsibly* can change the world.

Goals of this course

Our goal is to prepare students to effectively apply machine learning methods to problems that might arise in "the real world" -- in industry, medicine, government, education, and beyond.

Gain skills and *understanding* for a future as:

- Developer using ML “out-of-the-box”
- ML methods researcher

After taking this course, you will be able to:

- *Think systematically and ethically*
  - Compare/contrast each method's strengths & limitations
  - "Can ML solve this problem?"
  - "***Should*** ML solve this problem?"

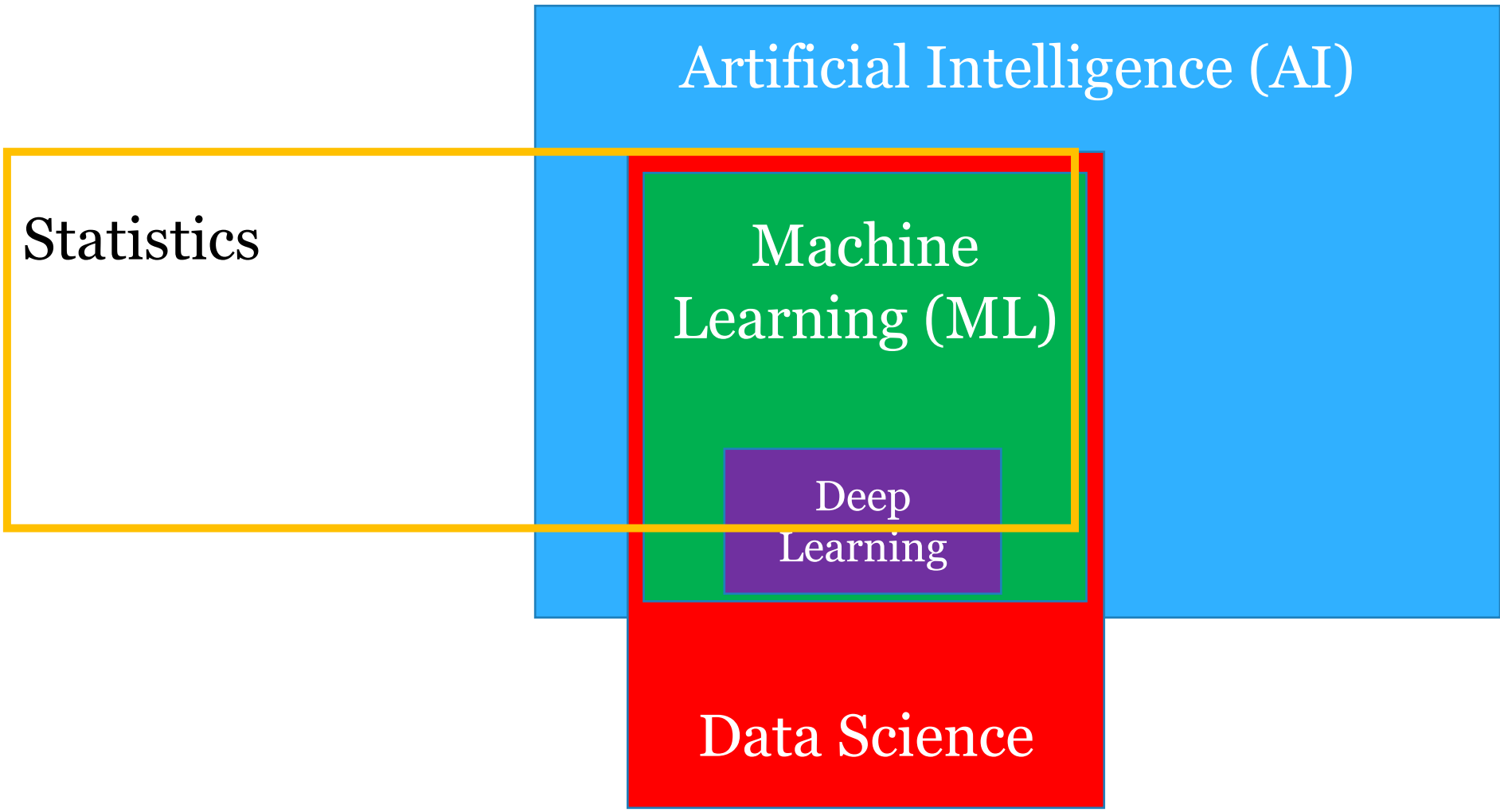
- *Deploy and debug rapidly on real problems*
  - Hands-on experience with open-source libraries
  - Address issues in "real-world" data analysis
    - Preprocessing, Numerical issues, convergence issues, class imbalance, etc.

- *Evaluate carefully and honestly*
  - Design experiments to assess generalization to never-before-seen data
  - Select task-appropriate performance metrics
  - Report confidence or uncertainty in performance numbers

- *Communicate insightfully and reproducibly*
  - Surface key insights via figures, tables, and text in a written report
  - Provide details for a peer to repeat your analysis and draw same conclusions

Q: What is Machine Learning?

Venn Diagram of Knowledge



Artificial Intelligence (AI)

Study of "intelligent" systems, with many parts:

logic, planning, search, probabilistic reasoning, **learning from experience,** interacting with other agents, etc.

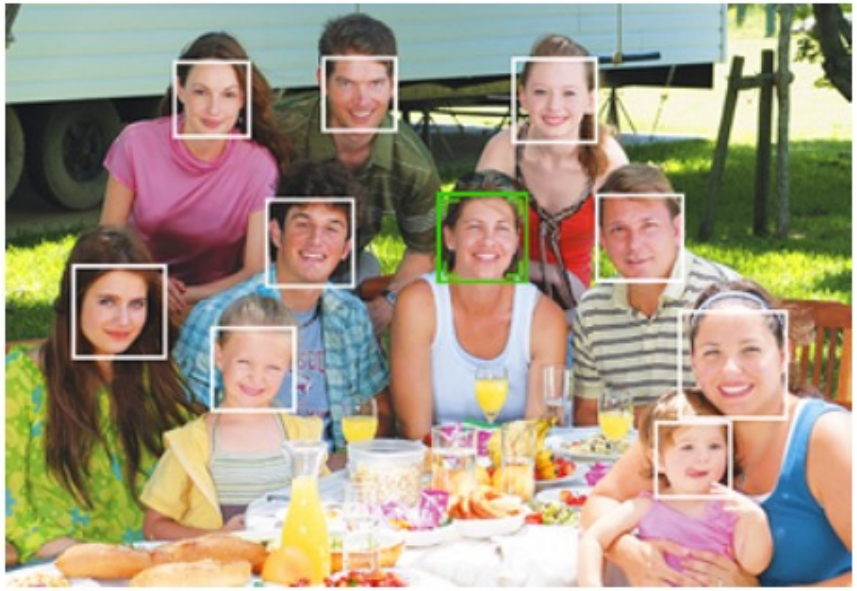
Alpha Go
Computer that can beat best human players of the game of "Go" (harder than chess)

DARPA Grand Challenge
Autonomous vehicles can navigate a real-world course without humans at the wheel

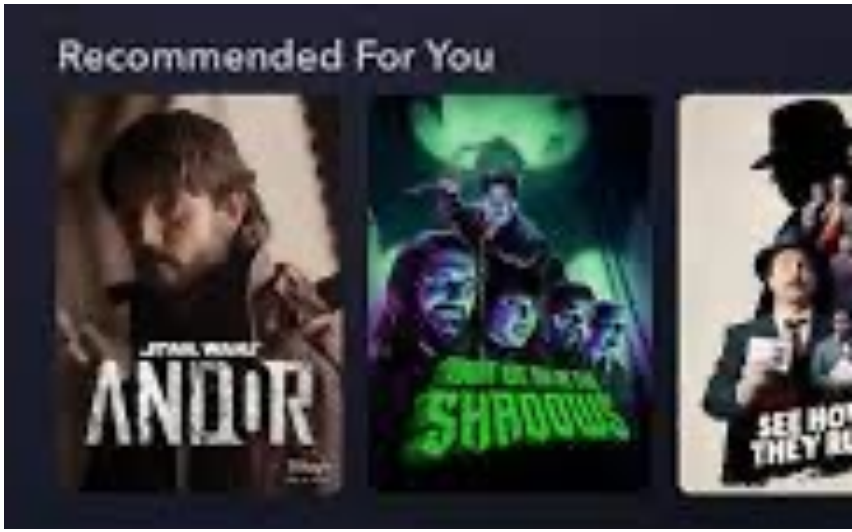
Machine Learning (ML)

Study of computer programs that **learn from experience/data** to perform a task

- Output: a *prediction, decision, or summary*

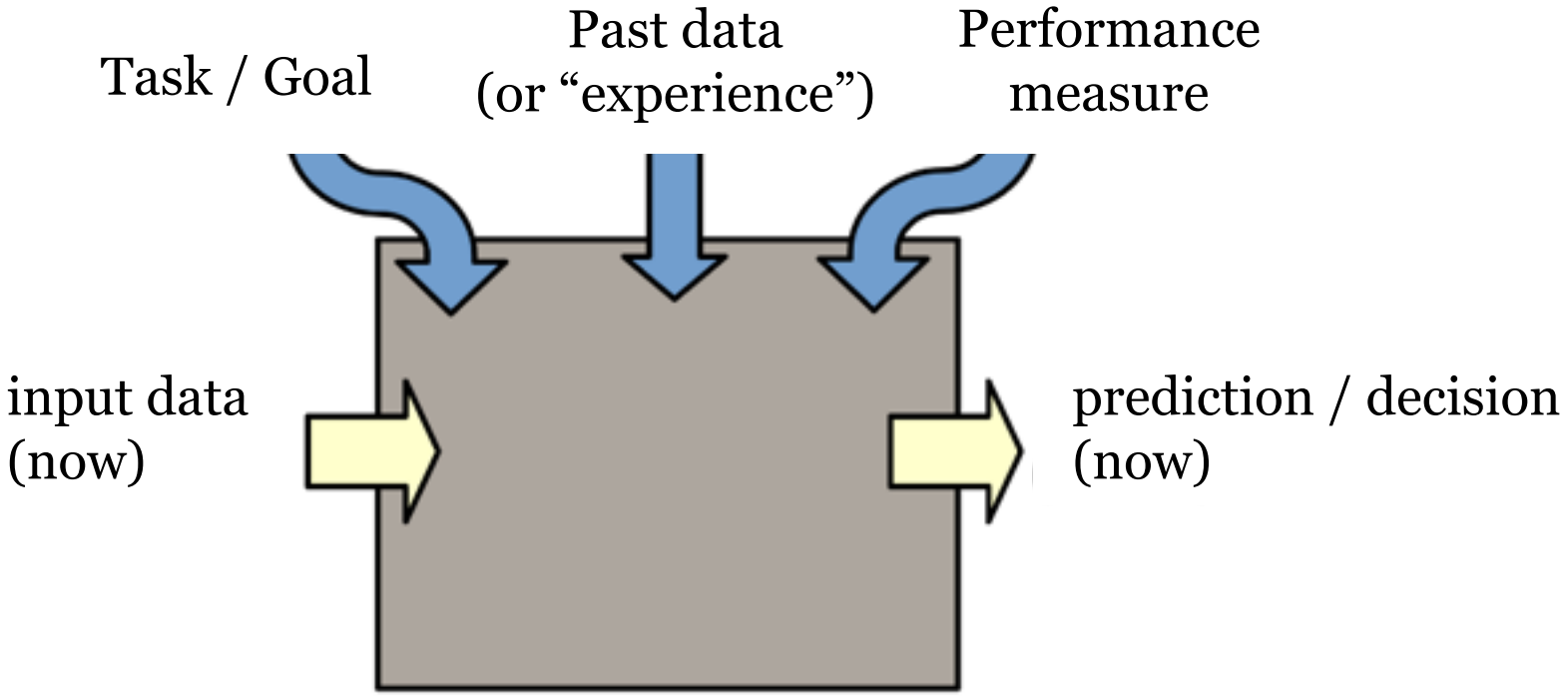


Face Detection
Predict location of human faces in natural images

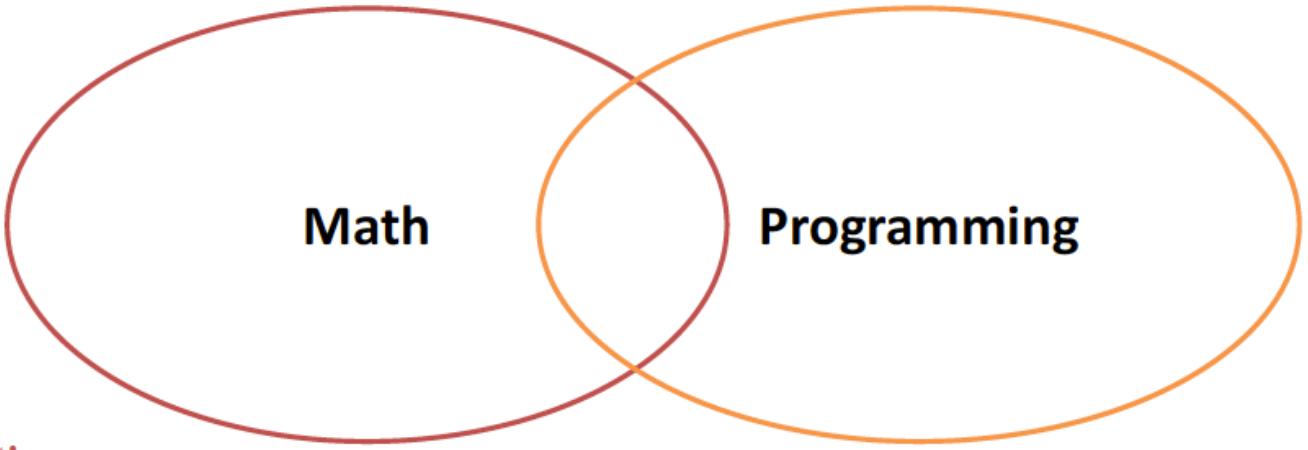


Movie Recommendation
Predict what to watch next

The Machine Learning Process



Q: What concepts will we learn?



Statistics,
Probability,
Linear Algebra,
Optimization

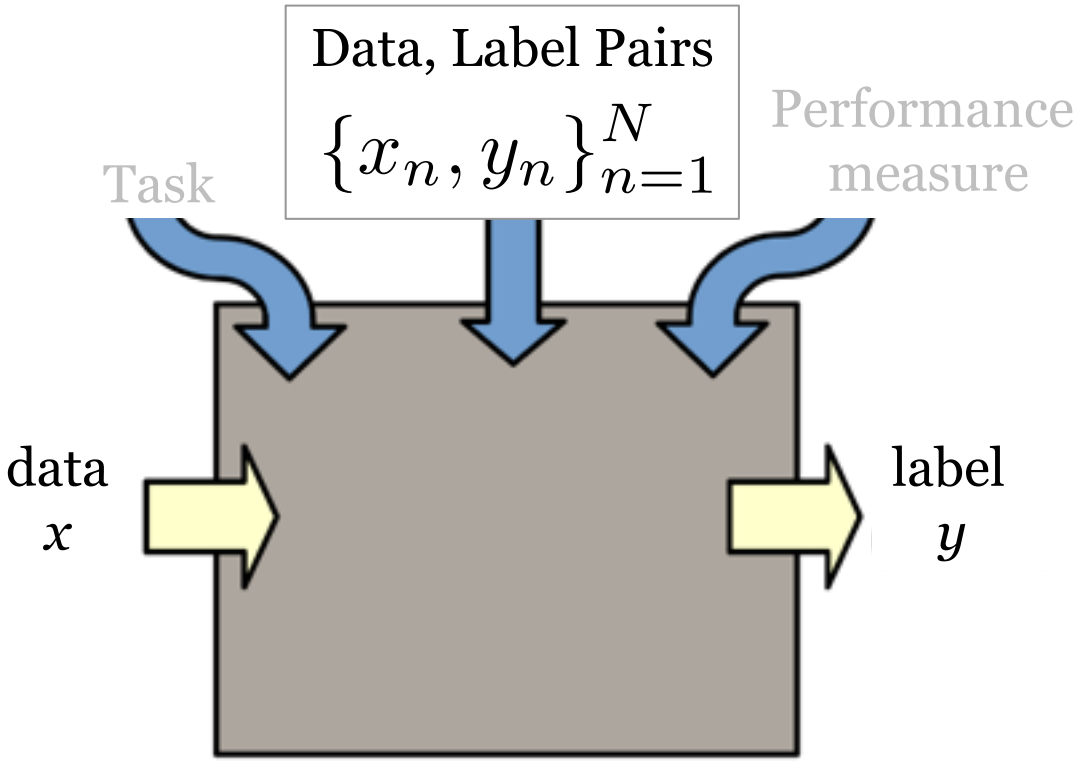
Data Structures,
Algorithms,
Computational Complexity,
Data Management

What will we learn?

Supervised Learning

Unsupervised Learning

Reinforcement Learning



Task: Regression

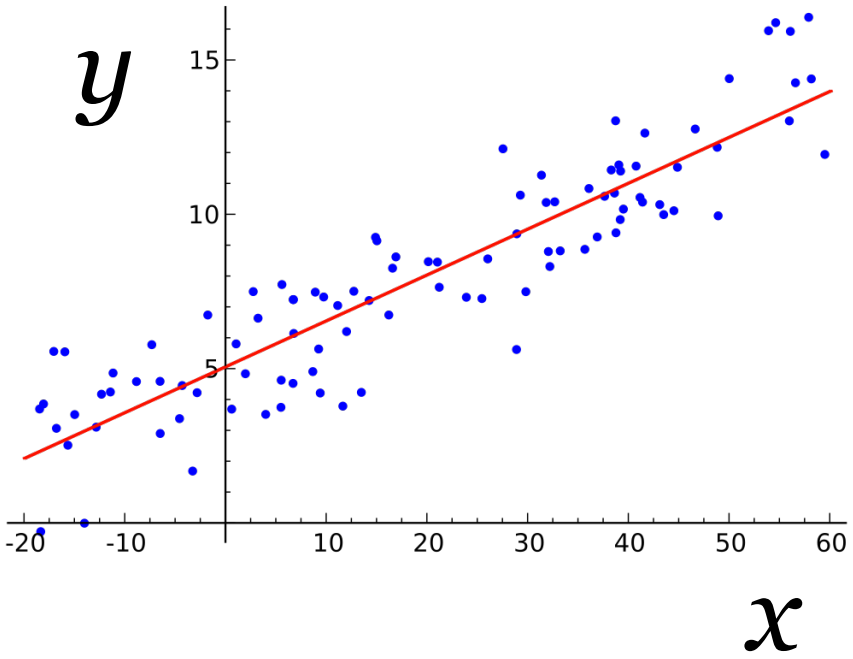
Supervised Learning

regression

Unsupervised Learning

Reinforcement Learning

y is a continuous variable
e.g. sales in $$



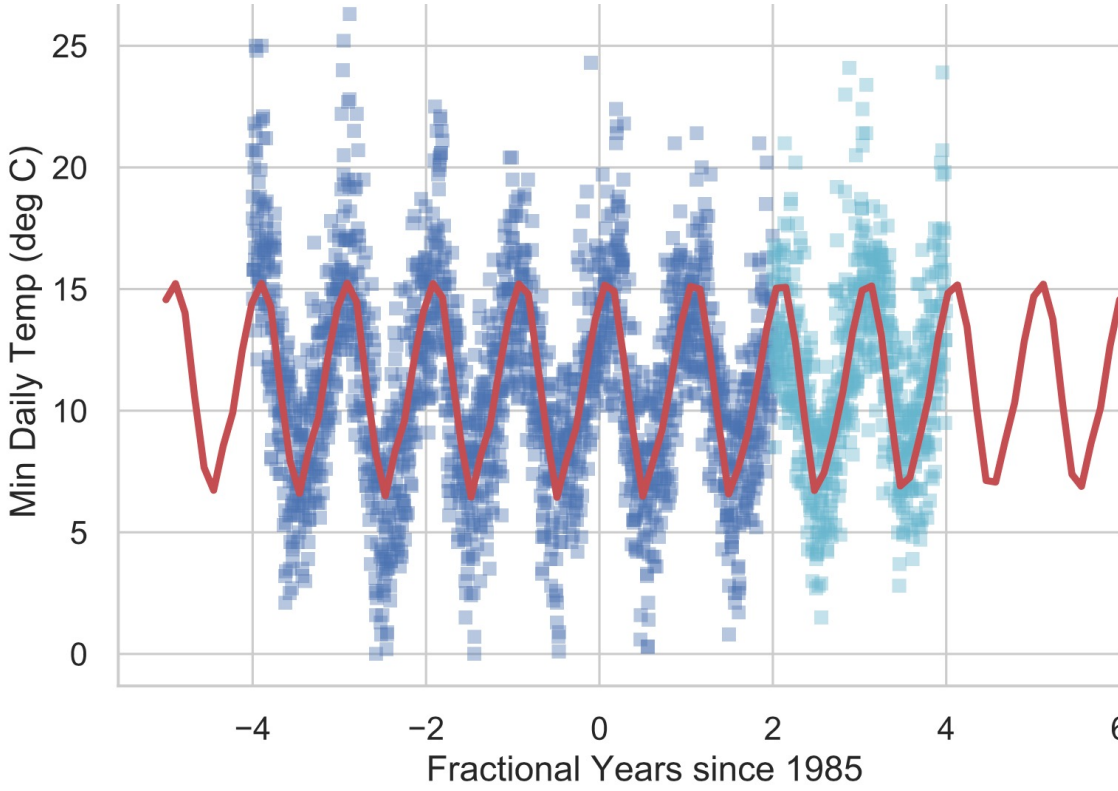
Regression: Temp over Time

Supervised Learning

regression

Unsupervised Learning

Reinforcement Learning



Task: Classification

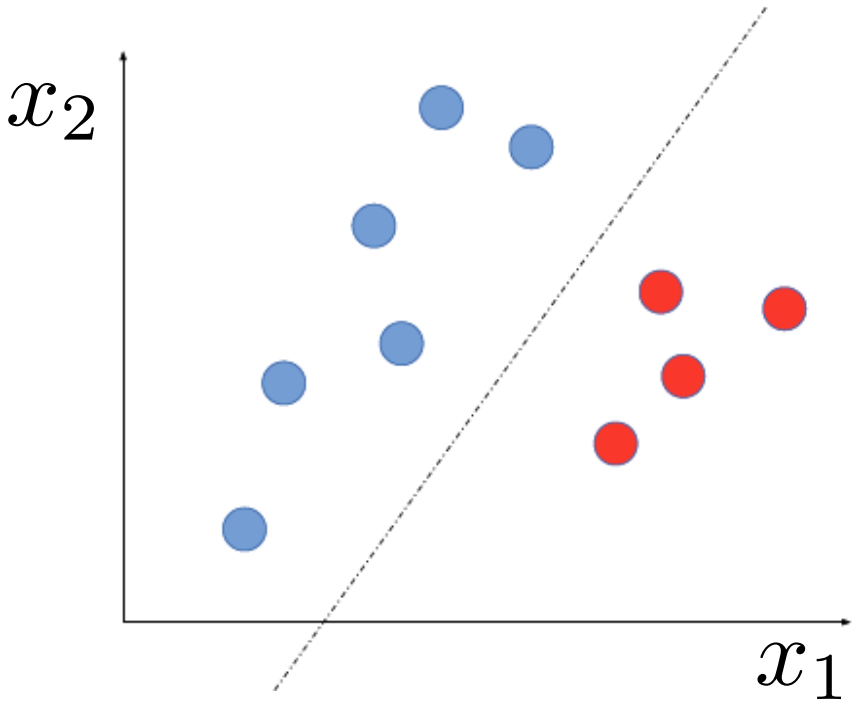
Supervised Learning

classification

Unsupervised Learning

Reinforcement Learning

y is a discrete variable (red or blue)



Classify: Pos or Neg review?

Supervised Learning

classification

Unsupervised Learning

Reinforcement Learning

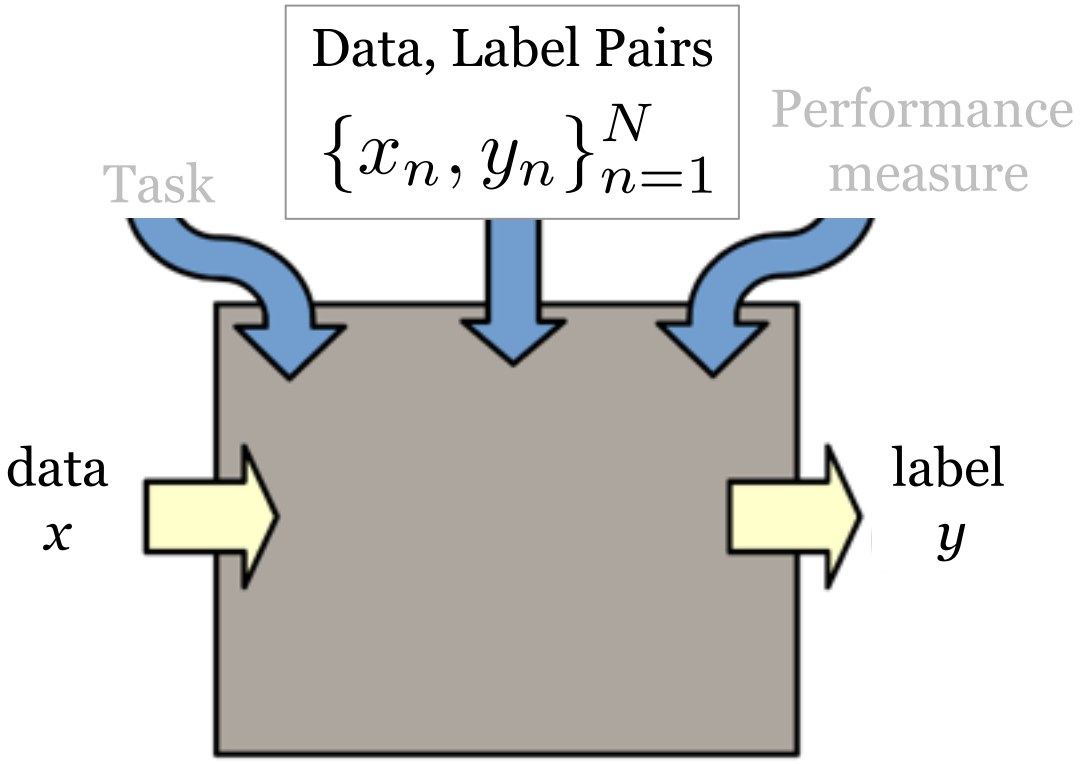
- Food was so gooodd
- I could eat their bruschetta all day it is devine.
- not sure how long we stood there but it was long enough for me to begin to feel awkwardly

What will we learn?

Supervised Learning

Unsupervised Learning

Reinforcement Learning

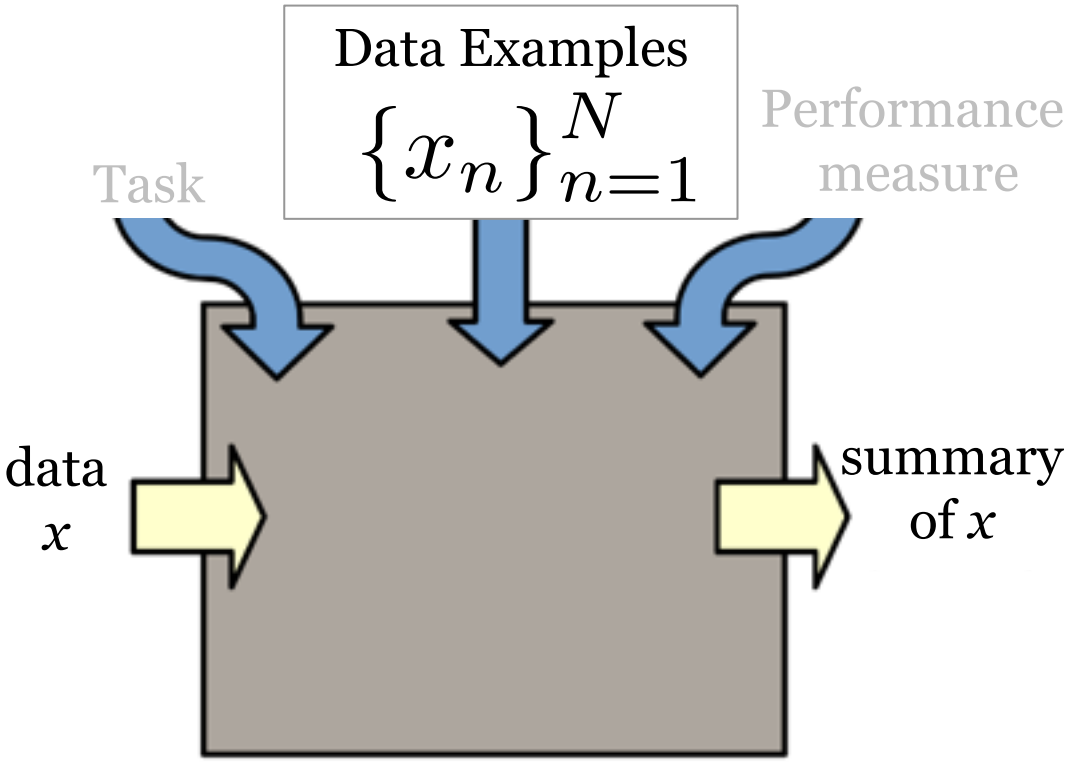


What will we learn?

Supervised Learning

Unsupervised Learning

Reinforcement Learning



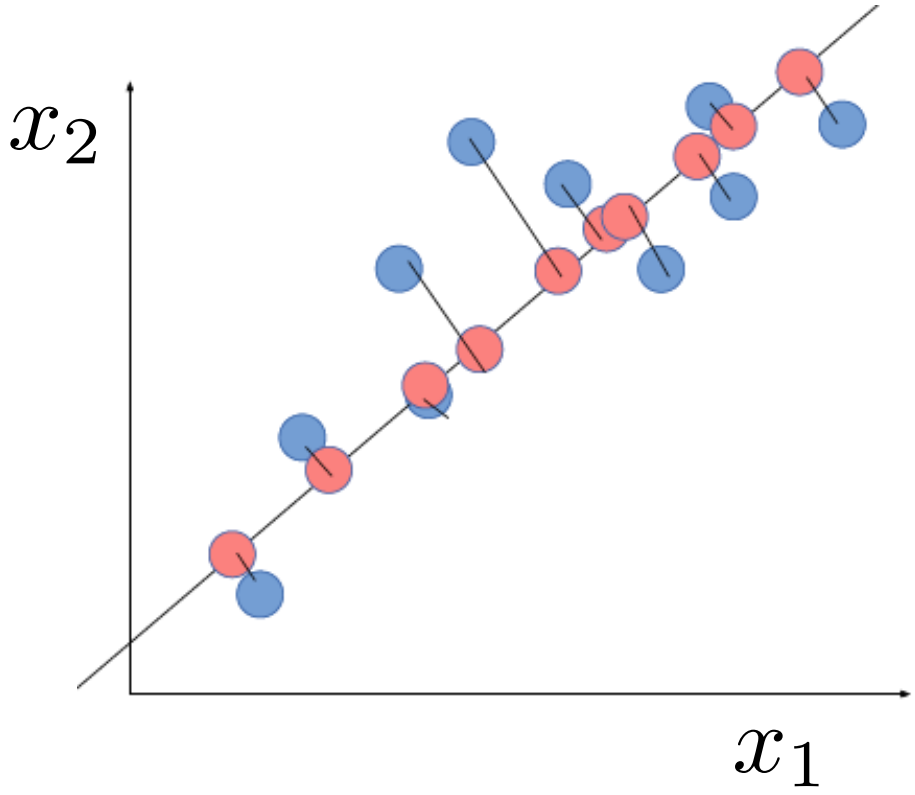
Task: Embedding

Supervised Learning

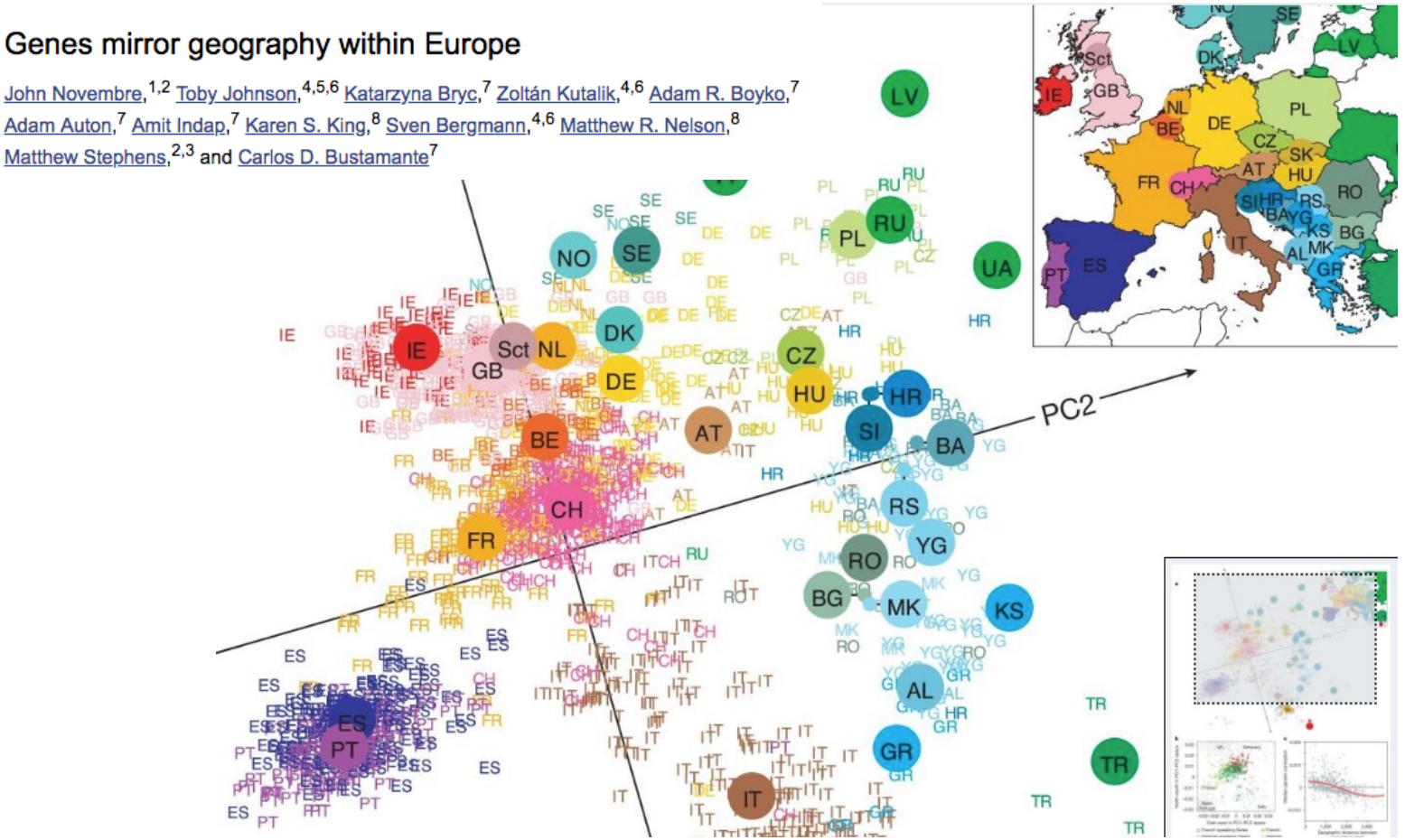
Unsupervised Learning

embedding

Reinforcement Learning



Example: Genes vs. geography



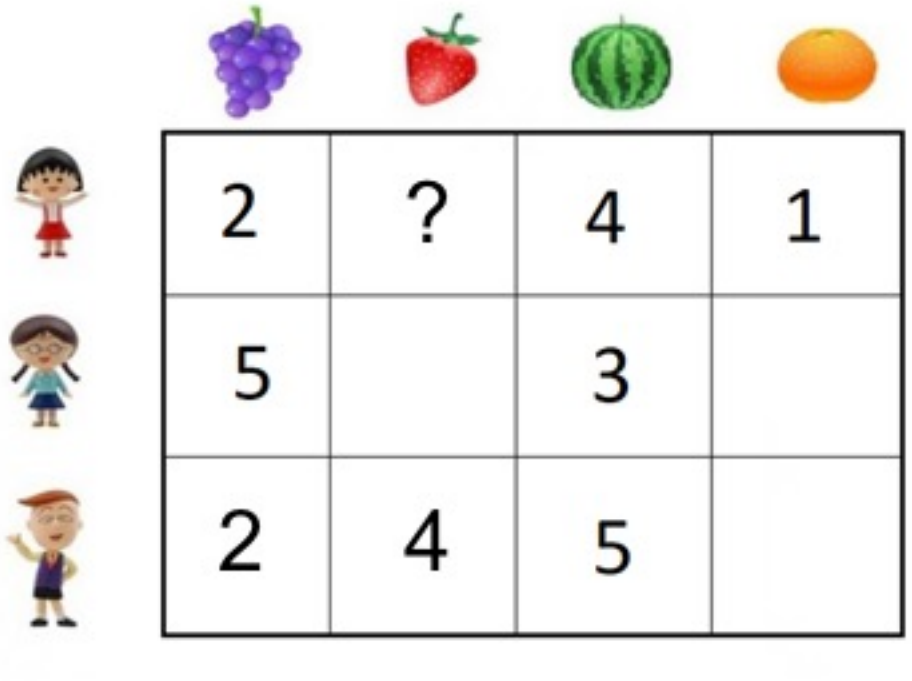
Task: Recommendation

Supervised Learning

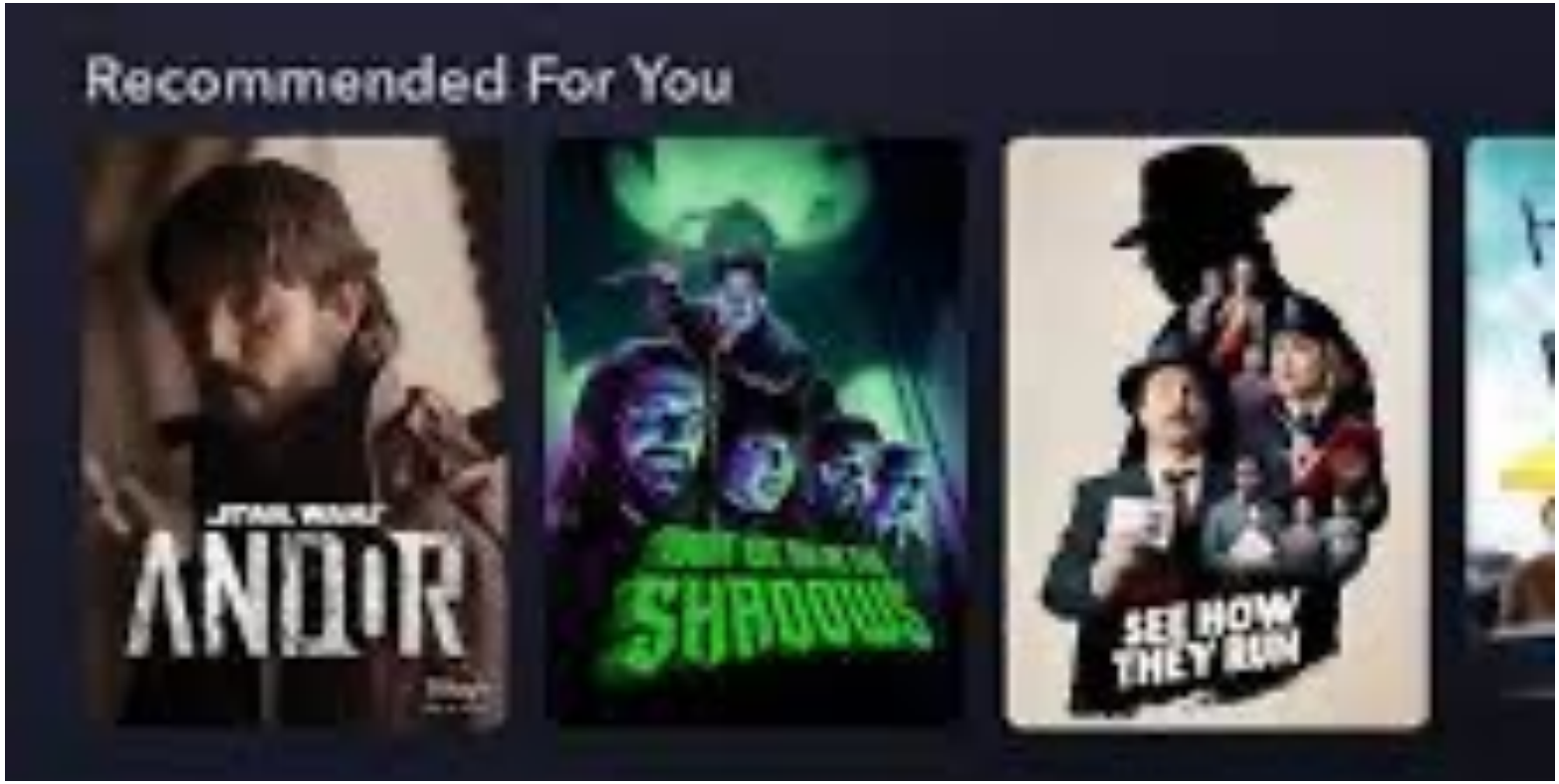
recommendation

rvised Learning

Reinforcement Learning



Recommendation Example



What **won't** we cover?

- Reinforcement learning
- Clustering
- Graphical models

- Active learning
- Transfer learning
- Semi-supervised learning

- Learning theory
- ..... lots more