Intro to Machine Learning (CS436/CS580L)

Lecture 1: Introduction

Xi Peng, Fall 2018

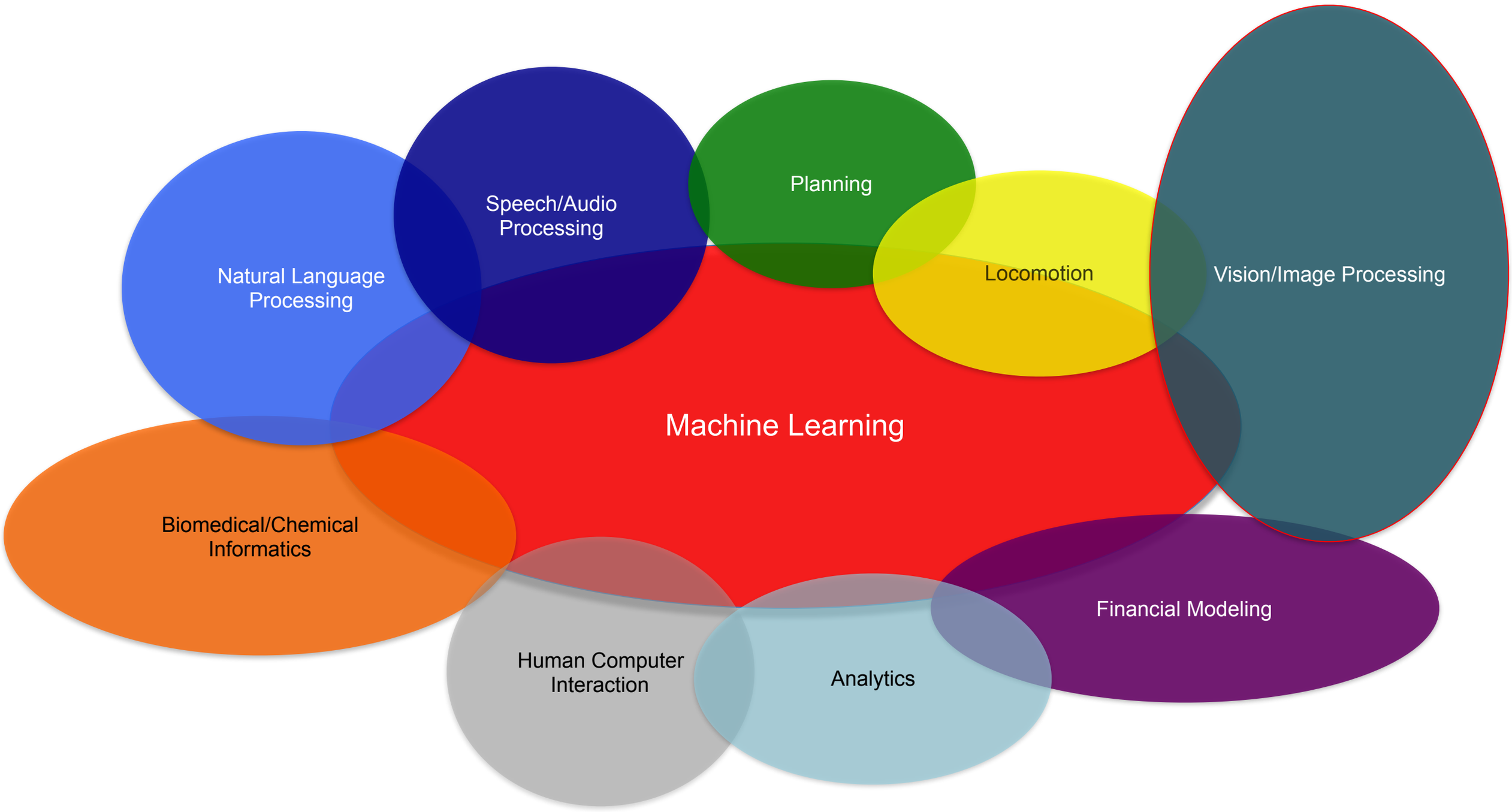
Thanks to Tom Mitchell, Andrew Ng, Feifei Li, Ben Taskar, Carlos Guestrin, Eric Xing, Hal Daume III, Jerry Zhu, Tina Eliassi-Rad, Vladimir Pavlovic, and Chao Chen for some slides & teaching material.

This Class

- ML in CS
- ML v.s. AI
- ML v.s. Statistics
- What is ML?
- Different type of ML:
  - Supervised learning
  - Unsupervised learning
  - Reinforcement learning

EASY

Machine Learning in Computer Science



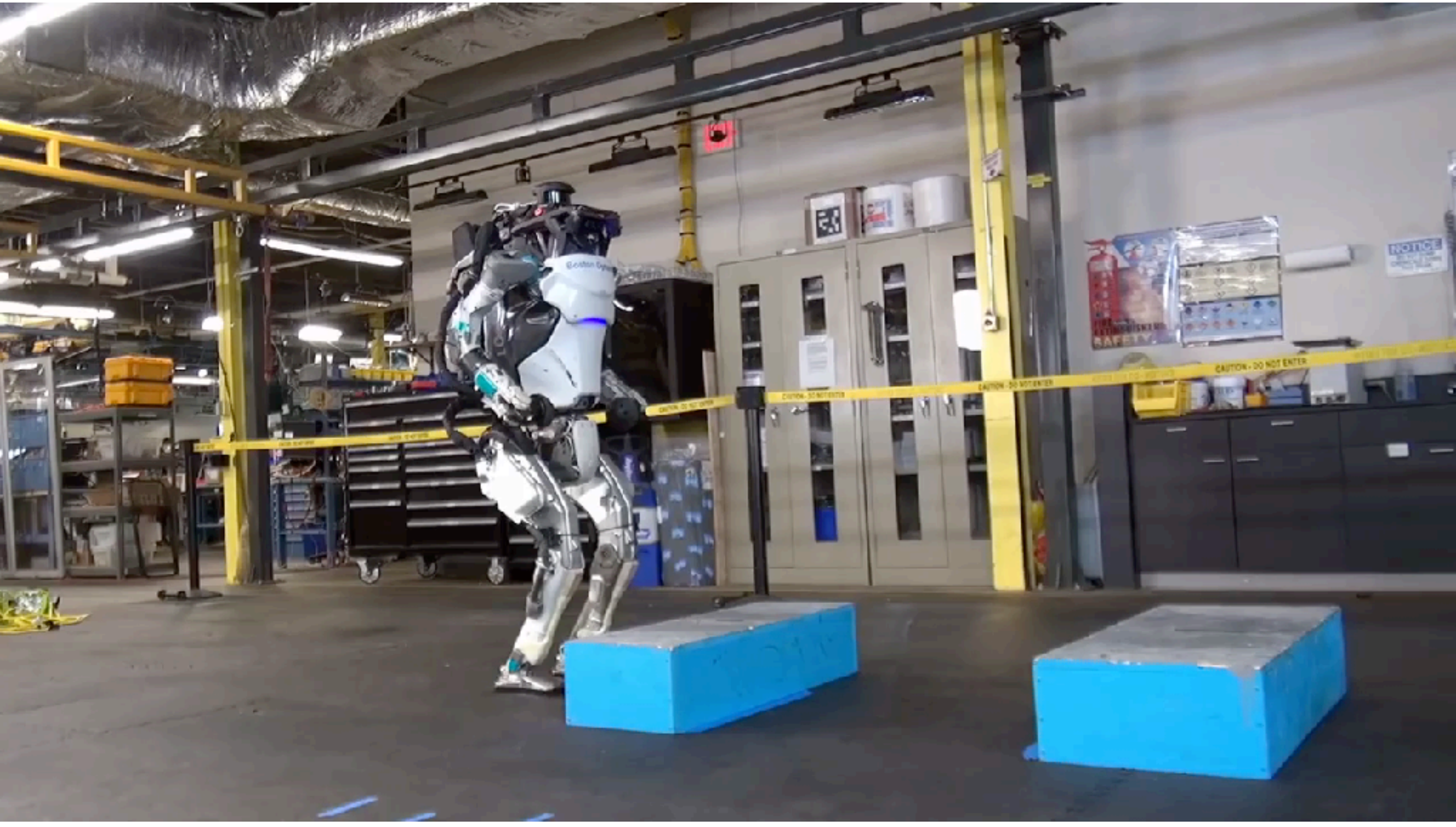
Google

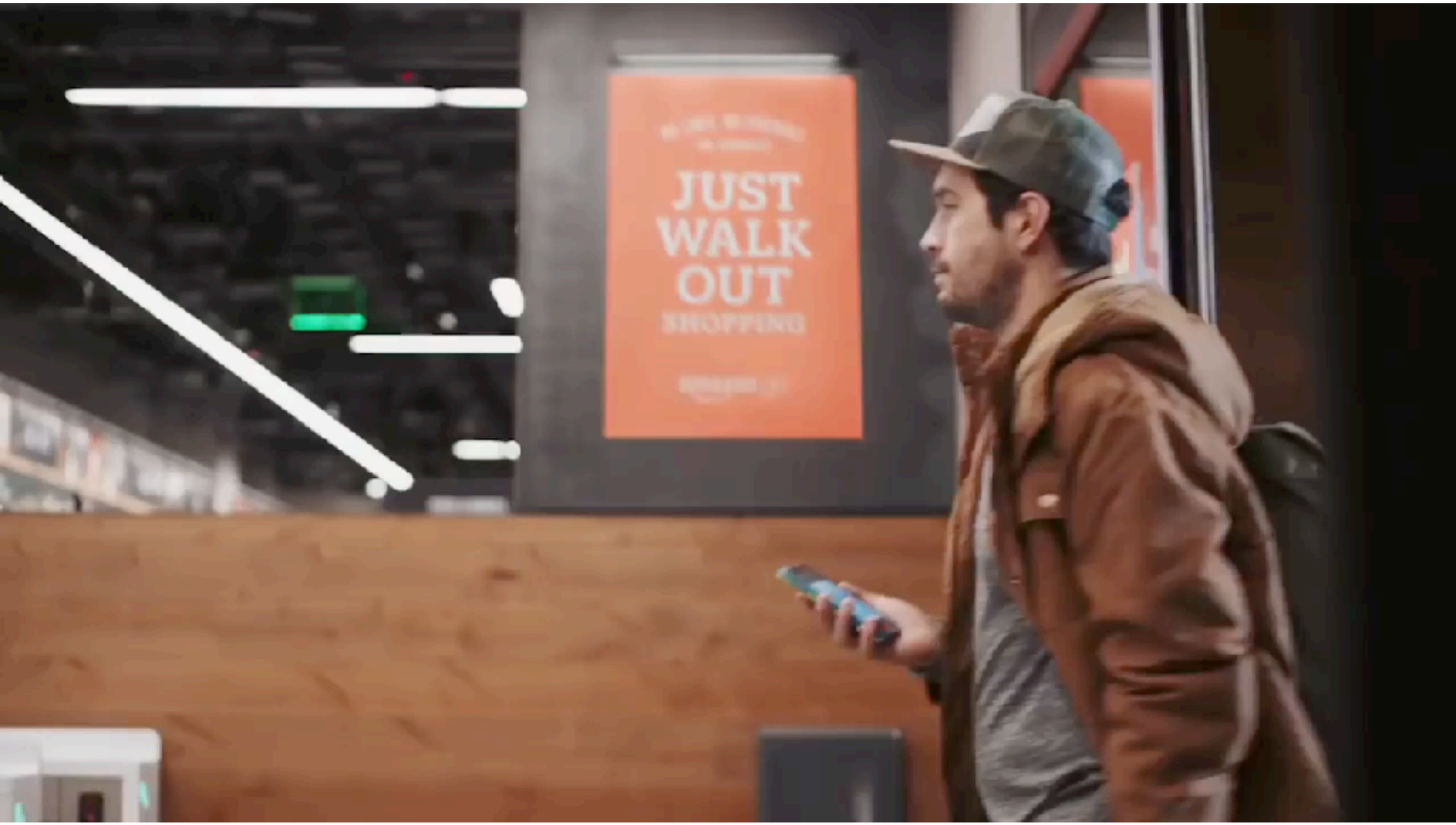
machi

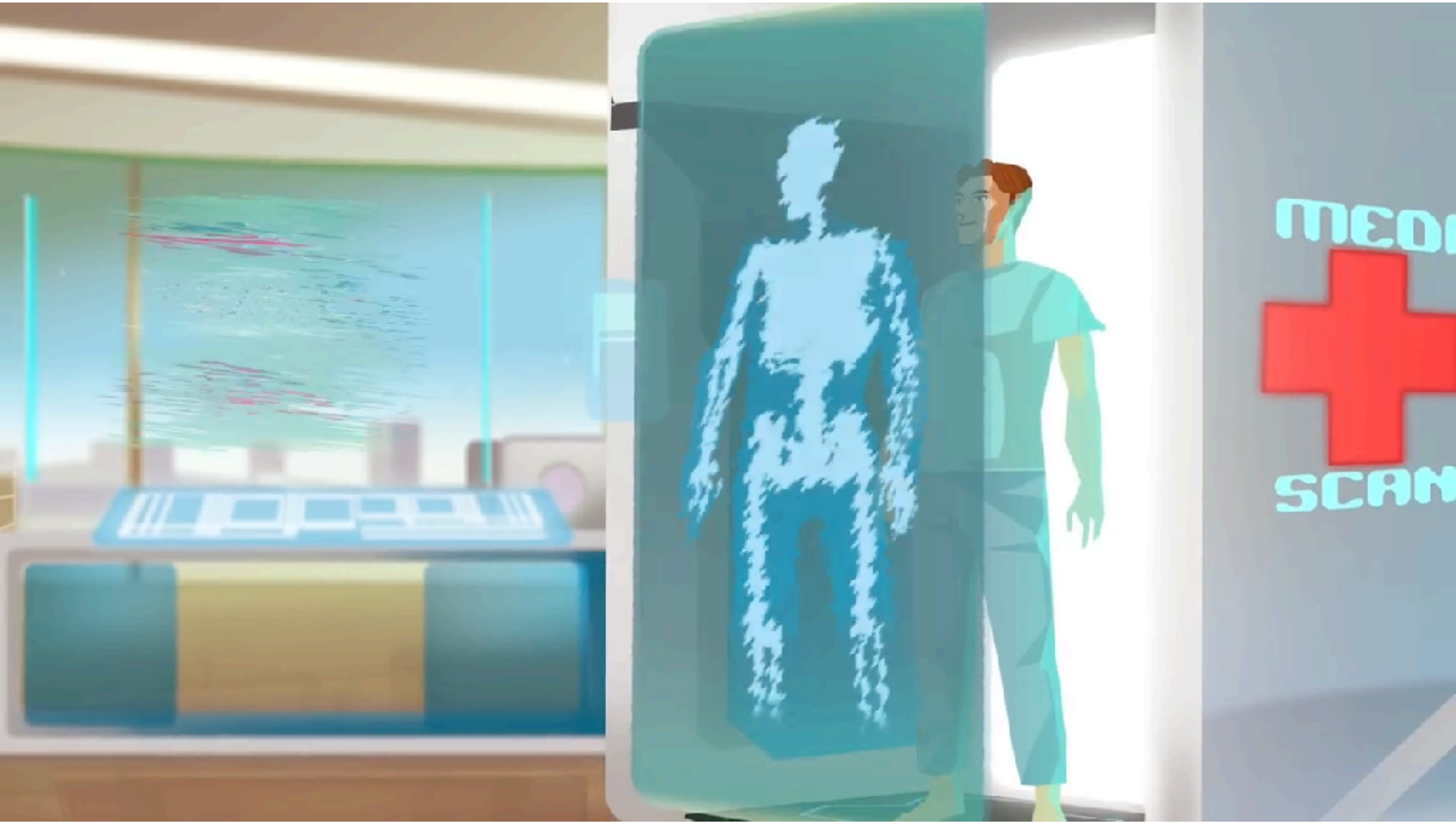
machine learning Remove
machine-assisted curation Remove
machine learning interview questions
machine gun kelly
machiavelli
machine learning algorithms
machine
machine learning course
machiavellian
machinery trader

Google Search I'm Feeling Lucky

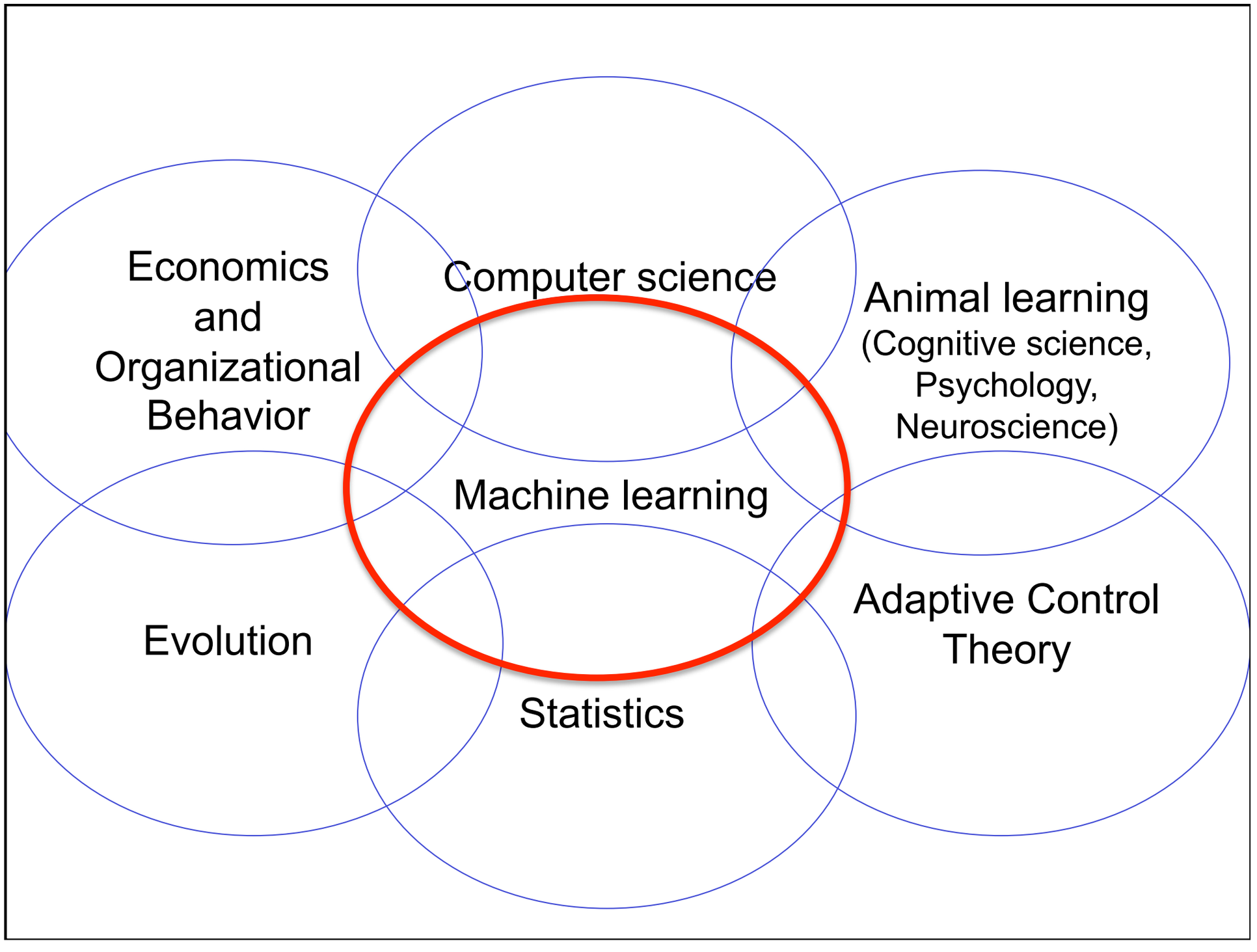
Report inappropriate predictions


CAUTION - DO NOT ENTER
NOTICE


JUST
WALK
OUT
SHOPPING


SCAN

Machine learning grow out of AI



Evolution from AI to Deep Learning

ARTIFICIAL INTELLIGENCE

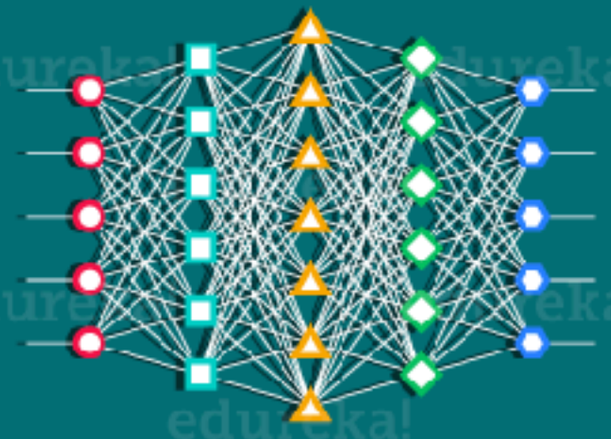
Engineering of making Intelligent Machines and Programs

MACHINE LEARNING

Ability to learn without being explicitly programmed

DEEP LEARNING

Learning based on Deep Neural Network



1950's | 1960's | 1970's | 1980's | 1990's | 2000's | 2006's | 2010's | 2012's | 2017's

CS, Stat, ML Questions

- Computer Science (CS)
  - How can we build machines that solve problems, and which problems are inherently tractable/intractable?
- Statistics (Stat)
  - What can be inferred from data plus a set of modeling assumptions, with what reliability?
- Machine Learning (ML)
  - How can we build computer systems that automatically improve with experience, and what are the fundamental laws that govern all learning processes?

What is Machine Learning (ML)

- Statistics + Algorithms
  - Have data
    - [sqft, # bdrms, # baths, age of house; *price of house*]
  - Select model
    - Linear regression
  - Select algorithm to learn parameters of model
    - Gradient descent
- Fancy Function Fitting
  - $\hat{f}(x) = y$

How is machine learning different from statistics?

- Statistics cares about the model,
  ML cares about predictions

- Statistics cares about model fit,
  ML cares about generalization

- Statistics tries to explain the world,
  ML tries to predict the future

What is Machine Learning (ML)

- Arthur Samuel (1959)
  - Field of study that gives computers the ability to learn without being explicitly programmed.
- Tom Mitchell (1997)
  - Well-posed Learning Problem (*T*, *E*, *P*):
    A computer program is said to **learn** from experience E w.r.t. some task T and some performance measure P, if its performance on T, as measured by P, **improves** with experience E.

Example 1

- Suppose your email program watches which emails you do or do not mark as spam, and based on that learns how to better filter spam.
- Task T:
  - Classifying emails as spam or not spam.
- Experience E:
  - Watching you label emails as spam or not spam.
- Performance P:
  - The number (or fraction) of emails correctly classified as spam/not spam.

Example 2

- Suppose we feed a learning algorithm a lot of historical weather data, and have it learn to predict weather.
- Task T:
  - Weather prediction
- Experience E:
  - The process of the algorithm examining a large amount of historical weather data.
- Performance P:
  - The probability of it correctly predicting a future date's weather.

Types of Machine Learning

- Supervised Learning (SL)
- Unsupervised Learning (UL)
- Reinforcement Learning (RL)

Types of Machine Learning

Supervised Learning

- Observe *n* training examples $D=\{x_i, y_i\}$
- Learn a function, *h*, mapping any input *x to output y*: $h(x) \approx y$

- Classification: predict a small number of discrete-valued outputs
- Regression: predict a continuous-valued output
- Ranking
- …

So far… Supervised Learning

Data: (x, y)
x is data, y is label

Goal: Learn a *function* to map x -> y

Examples: Classification, regression, object detection, semantic segmentation, image captioning, etc.

Cat

Classification

Example 3

- Suppose you are working on weather prediction, and you would like to predict whether or not it will be raining at 5pm tomorrow.
- You want to use a learning algorithm for this.
- Would you treat this as a classification or a regression problem?
  - Classification; trying to predict a small number of discrete-valued outputs

Example 3b

- Suppose you are working on weather prediction, and your weather station makes one of three predictions for each day's weather: Sunny, Cloudy or Rainy. You'd like to use a learning algorithm to predict tomorrow's weather. Would you treat this as a classification or a regression problem?
  - Classification, trying to predict one of a small number of discrete-valued outputs, such as whether it is Sunny, Cloudy, or Rainy.

Example 4

- Suppose you are working on stock market prediction.
- Typically tens of millions of shares of Microsoft stock are traded (i.e., bought/sold) each day.
- You would like to predict the number of Microsoft shares that will be traded tomorrow.
- Would you treat this as a classification or a regression problem?
  - Regression; trying to predict a continuous-valued output

Example 5

- Suppose you are working on stock market prediction.
- You would like to predict whether the US Dollar will go up against the Euro tomorrow (i.e., whether a dollar will be worth more euros tomorrow than it is worth today).
- Would you treat this as a classification or a regression problem?
  - Classification, trying to predict if the US Dollar goes up or that it does not.

(Supervised + *X*) Learning

Active learning	=	Supervised + oracle queries
Semi-supervised learning	=	Supervised + unlabeled data
Multi-task & transfer learning	=	Supervised + data from a related task(s)
Domain adaptation	=	Supervised + data from a source domain(s)
Multi-view learning	=	Supervised + alternative data views
Learning in a bandit setting	=	Supervised + limited feedback
Multiple-instance learning	=	Supervised + relaxed label granularity
Learning with expert knowledge	=	Supervised + side information
Weakly labeled data	=	Supervised + free (possibly inaccurate) labels
Human computation & crowd-sourcing	=	Supervised + free/inexpensive annotators
Learning on a budget	=	Supervised + misc. cost constraints
Learning under constraints	=	Supervised + constr. of low-dim structures

Types of Machine Learning

- Supervised Learning (SL)
- Unsupervised Learning (UL)
- Reinforcement Learning (RL)

Types of Machine Learning
Unsupervised Learning

- Observe *n* examples $D = \{x_i\}$
- Clustering: Find a well-separated partition of example $D=\{D_1, D_2, \ldots, D_k\}$
- Dimensionality reduction: Find a low-distortion, low-dimensional projection $y = P\,x$, where the number of dimensions of *y* is less than (sometimes much less than) the number of dimensions of *x*
- The goal is to replace the high-complexity description of with a lower-complexity one

So far… Unsupervised Learning

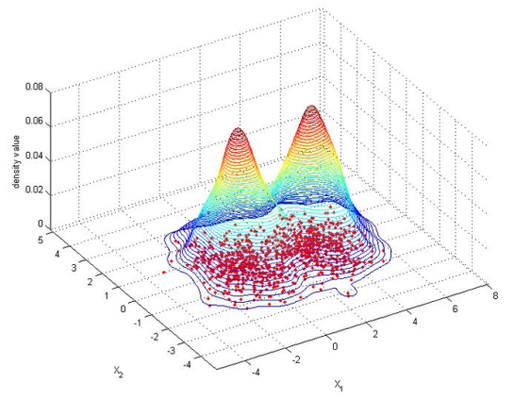
Data: x
Just data, no labels!

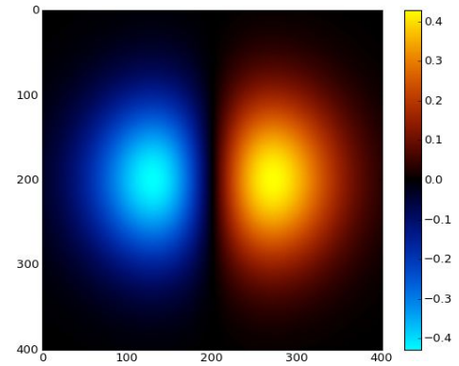
Goal: Learn some underlying hidden *structure* of the data

Examples: Clustering, dimensionality reduction, feature learning, density estimation, etc.

Figure copyright Ian Goodfellow, 2016. Reproduced with permission.

1-d density estimation





2-d density estimation

2-d density images left and right are CC0 public domain

Example 6
Supervised vs. Unsupervised Learning

- Given historical data of childrens' ages and heights, predict children's height as a function of their age.
  - Supervised, regression problem, learning to predict height from a training set
- Examine a large collection of emails that are known to be spam email, to discover if there are sub-types of spam mail.
  - Unsupervised learning, cluster spam mail into sub-types
- Given 50 articles written by male authors, and 50 articles written by female authors, learn to predict the gender of a new manuscript's author (when the identity of this author is unknown).
  - Supervised learning, classification, learn to predict gender from labeled data
- Take a collection of 1000 essays written on the US Economy, and find a way to automatically group these essays into a small number of groups of essays that are somehow "similar" or "related".
  - Unsupervised learning, similar to Google News

Example 6b
Supervised vs. Unsupervised Learning

- Given genetic (DNA) data from a person, predict the odds of him/her developing diabetes over the next 10 years.
  - Supervised learning, classification problem, where we can learn from a labeled dataset comprising different people's genetic data, and labels telling us if they had developed diabetes
- In farming, given data on crop yields over the last 50 years, learn to predict next year's crop yields.
  - Supervised learning problem, where we learn from historical data (labeled with historical crop yields) to predict future crop yields.

- Given a large dataset of medical records from patients suffering from heart disease, try to learn whether there might be different clusters of such patients for which we might tailor separate treatments.
  - Unsupervised learning
- In farming, given data on crop yields over the last 50 years, learn to predict next year's crop yields.
  - Supervised learning
- Given data on how 1000 medical patients respond to an experimental drug (such as effectiveness of the treatment, side effects, etc.), discover whether there are different categories or "types" of patients in terms of how they respond to the drug, and if so what these categories are.
  - Unsupervised learning
- Examine a web page, and classify whether the content on the web page should be considered "child friendly" (e.g., non-pornographic, etc.) or "adult."
  - Supervised learning

Types of Machine Learning

- Supervised Learning (SL)
- Unsupervised Learning (UL)
- Reinforcement Learning (RL)

So far… Supervised Learning

Data: (x, y)
x is data, y is label

Goal: Learn a *function* to map x -> y

Examples: Classification, regression, object detection, semantic segmentation, image captioning, etc.

Cat

Classification

So far… Unsupervised Learning

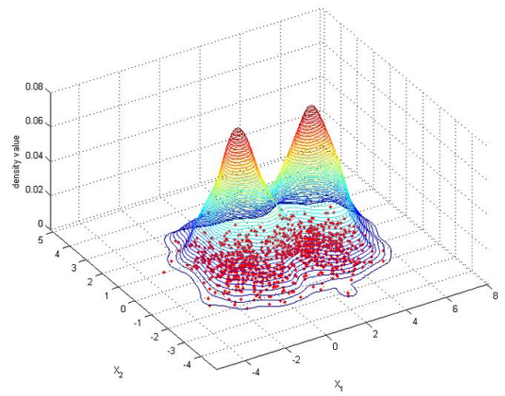
Data: x
Just data, no labels!

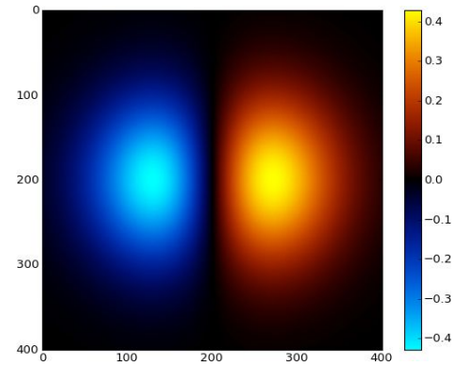
Goal: Learn some underlying hidden *structure* of the data

Examples: Clustering, dimensionality reduction, feature learning, density estimation, etc.

Figure copyright Ian Goodfellow, 2016. Reproduced with permission.

1-d density estimation



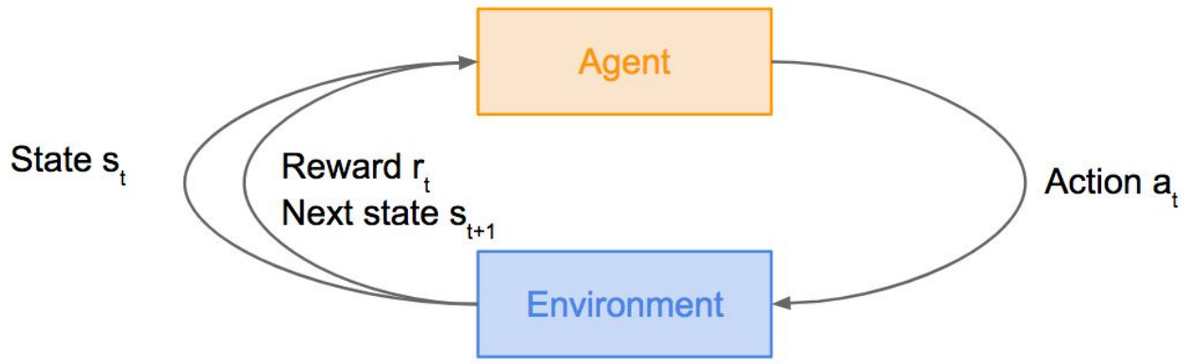


2-d density estimation

2-d density images left and right are CC0 public domain

Today: Reinforcement Learning

Problems involving an **agent** interacting with an **environment**, which provides numeric **reward** signals



Goal: Learn how to take actions in order to maximize reward

Types of Machine Learning

Reinforcement Learning

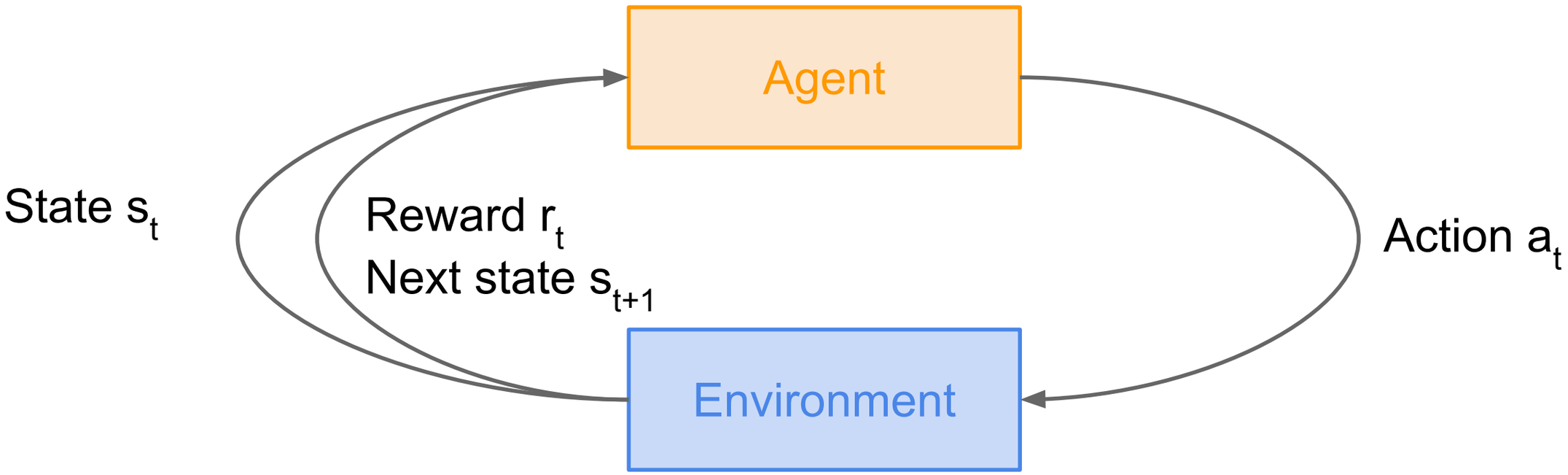
- Learn to act from delayed feedback
- Policy, π, maps system states to actions

$$\pi(s) = a$$

- Learning consists of
  - trying control policies $\{\pi_i\}$ on system
  - observing system states $\{S_i\}$
  - receiving rewards $\{R_i\}$

  to come up with a policy π with high expected reward

Reinforcement Learning

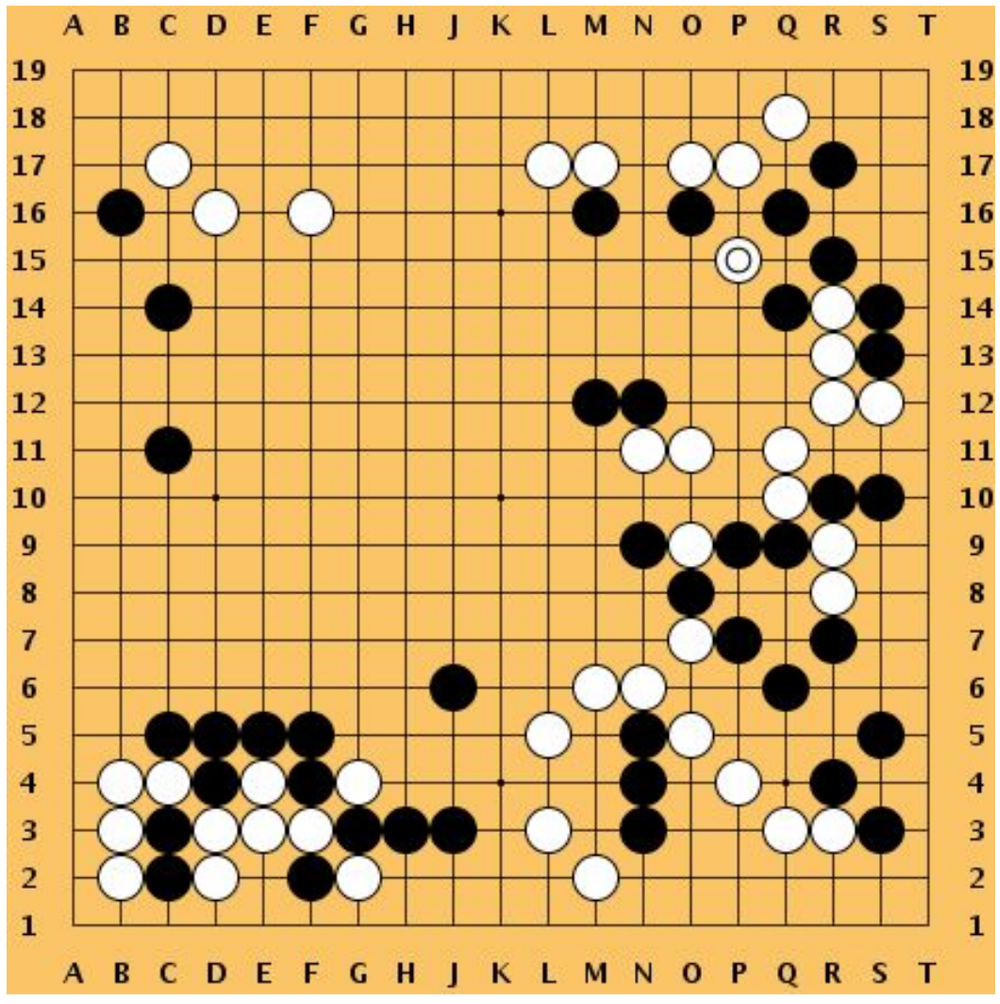


Atari Games

Objective: Complete the game with the highest score

State: Raw pixel inputs of the game state
Action: Game controls e.g. Left, Right, Up, Down
Reward: Score increase/decrease at each time step

Go



Objective: Win the game!

State: Position of all pieces
Action: Where to put the next piece down
Reward: 1 if win at the end of the game, 0 otherwise

This Class

- ML in CS
- ML v.s. AI
- ML v.s. Statistics
- What is ML?
- Different type of ML:
  - Supervised learning
  - Unsupervised learning
  - Reinforcement learning

EASY

Next Class

- Classification v.s. Regression
- Model representation
- Loss function

Median

To Do

- Read “Bishop”: Ch1 & Ch 2.
- Read “Statistics” and “Linear Algebra” resources.
- Pre-proposal of class project.