

Course Announcements

Due Sunday (11:59 PM):

- D7
- Q8
- Weekly Project Survey (*optional*)

Notes:

- iClicker Cloud attendance *should* now be accurate
- Discuss: EDA Checkpoint
- Reminder: Feedback for Sid (weekly project survey)

Machine Learning II & Ethics

Sid Joshi
UC San Diego

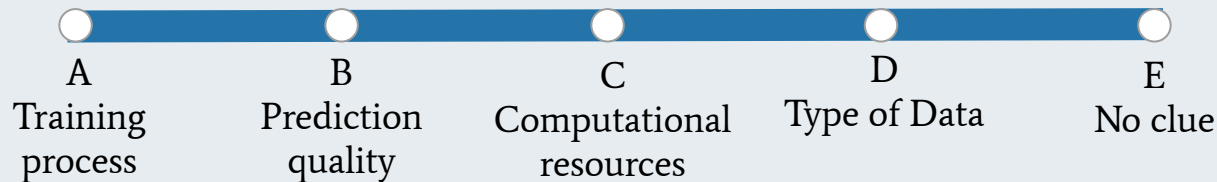


Department of Cognitive Science
s1joshi@ucsd.edu

Warm up

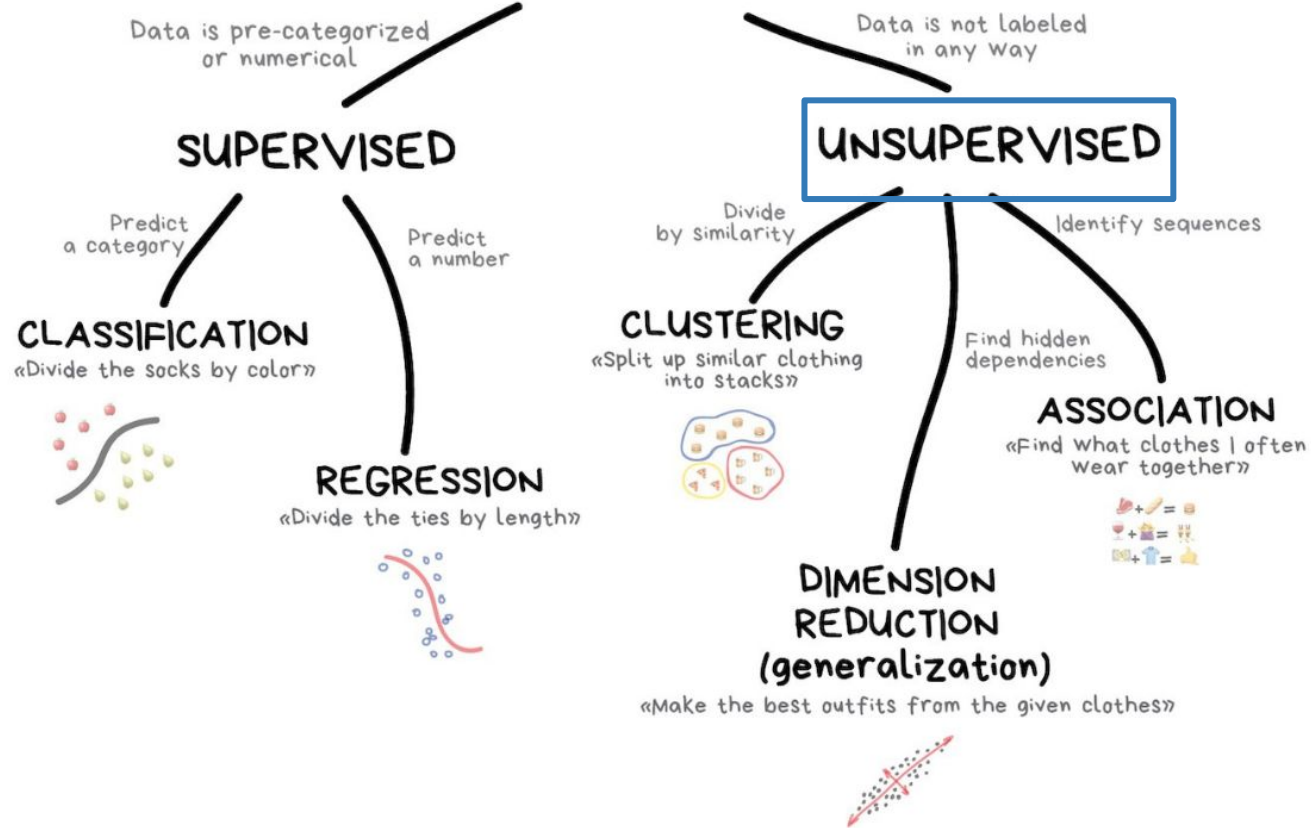


Where does the main difference lie between supervised and unsupervised learning?



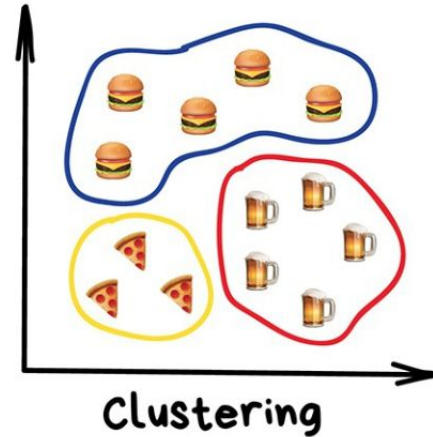
Unsupervised Machine Learning

CLASSICAL MACHINE LEARNING



Clustering

- Groups data points into clusters based on how **similar** they are to one another
- Similarity is usually defined by spatial relationships
- Used for:
 - Analyze and label new data
 - Anomaly detection
 - Merging points on a map
- Popular clustering models include:
 - K-means clustering
 - Gaussian Mixture
 - DBSCAN

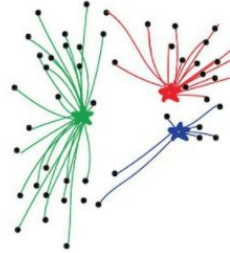


PUT KEBAB KIOSKS IN THE OPTIMAL WAY

(also illustrating the K-means method)



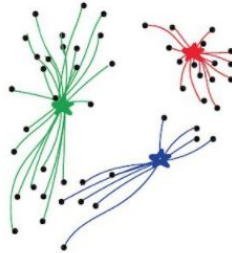
1. Put kebab kiosks in random places in city



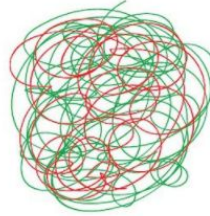
2. Watch how buyers choose the nearest one



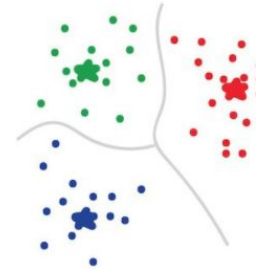
3. Move kiosks closer to the centers of their popularity



4. Watch and move again

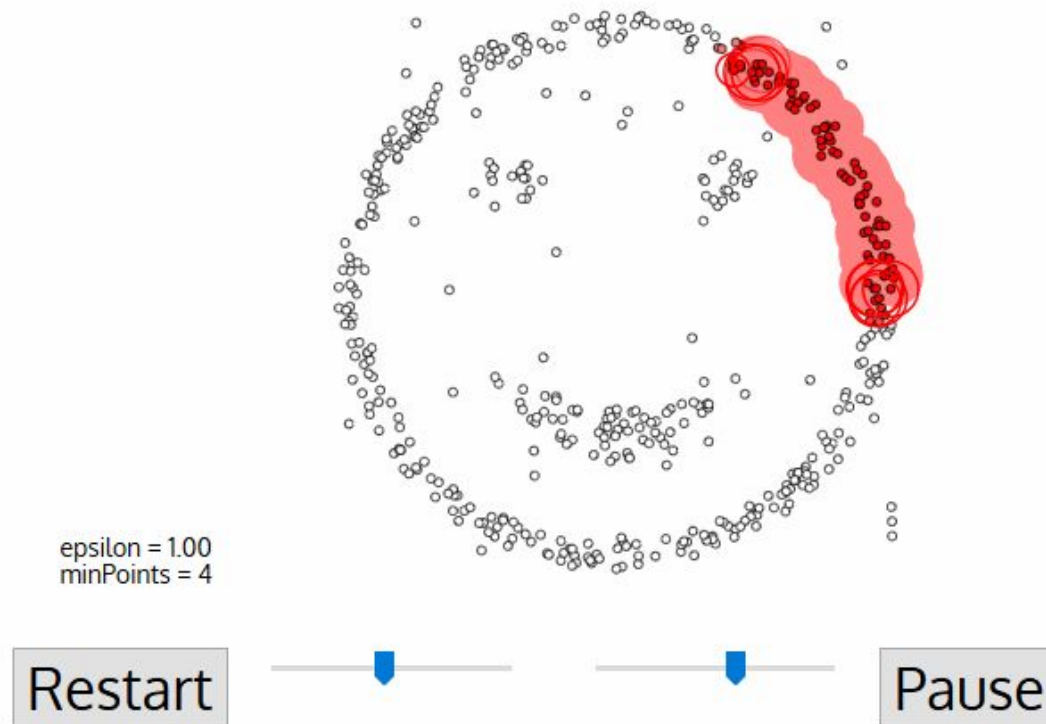


5. Repeat a million times



6. Done!
You're god of kebabs!

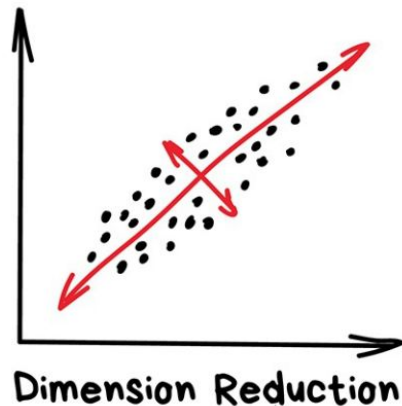
K-Means



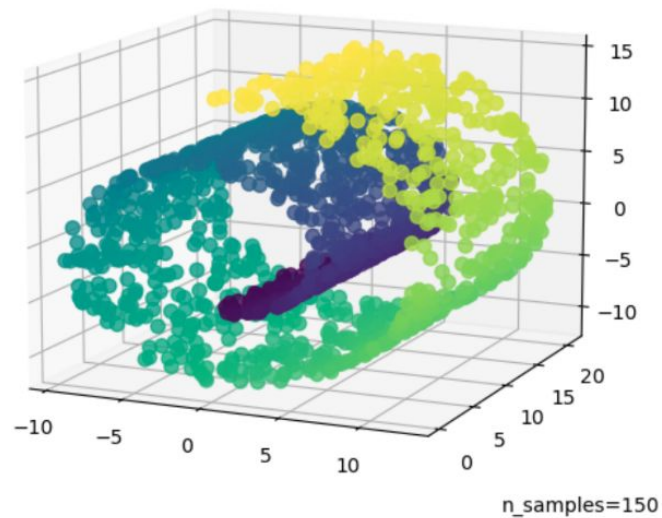
DBSCAN

Dimensionality Reduction

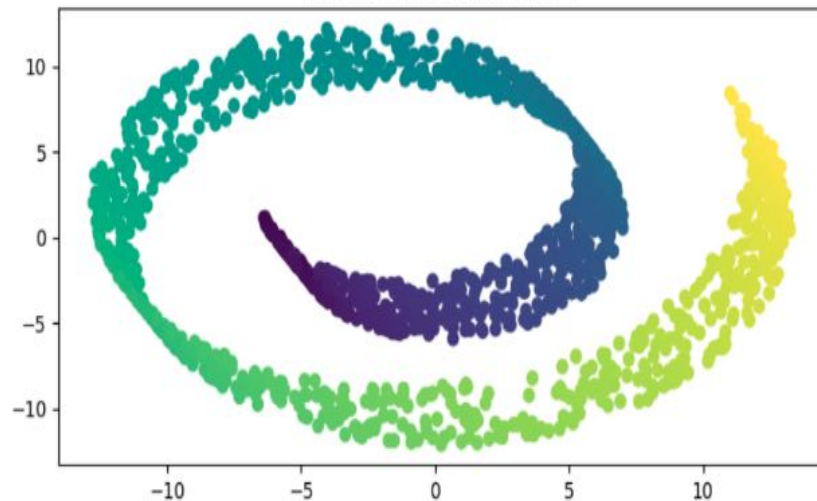
- Reduce high-dimensional data down to fewer dimensions
- Transforms data in a manner that preserves important, explanatory information
- Used for:
 - Fake image analysis
 - Data compression/feature selection
 - Genomics and bioinformatics
- Popular dimensionality reduction models include:
 - Principal component analysis (PCA)
 - t-SNE
 - UMAP



Swiss roll dataset

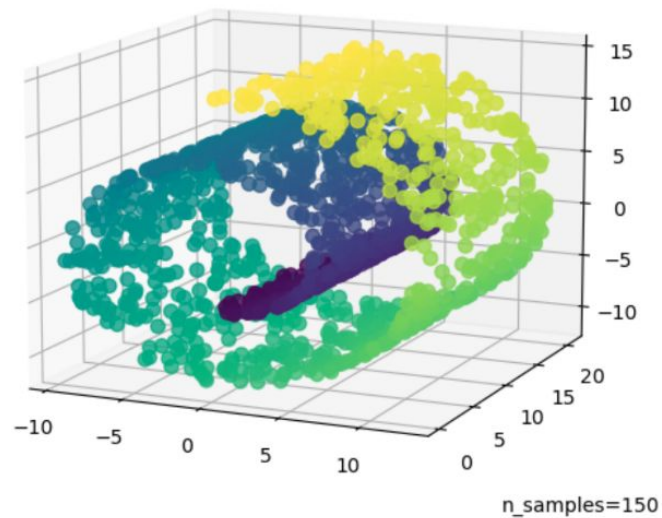


PCA of Swiss Roll Dataset

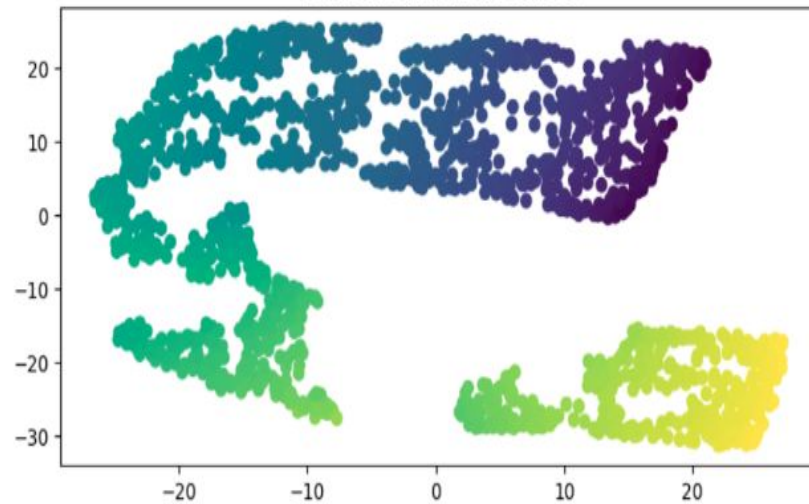


PCA

Swiss roll dataset



t-SNE of Swiss Roll Dataset



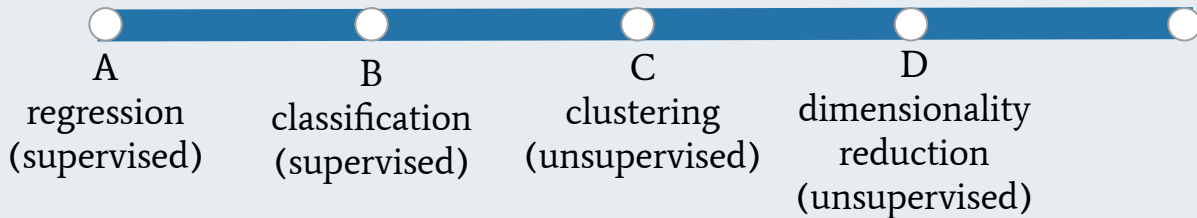
t-SNE

Prediction Approach



You have data on 10,000 socks and want to sort them based on color as either red, blue, green, or yellow.

What prediction approach would you use?

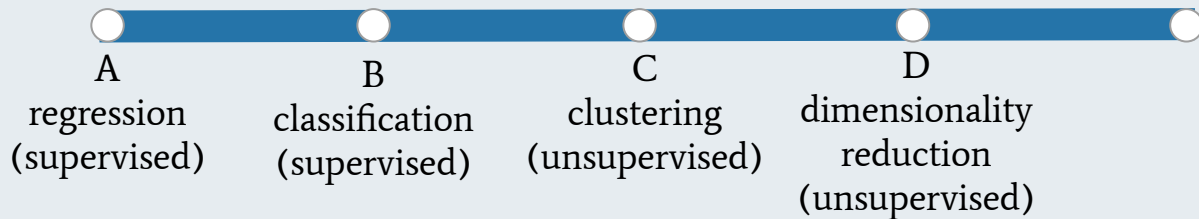


Prediction Approach



You have data on the gene expression levels of thousands of neurons from a new region in the brain. You want to gain new insight into expression patterns.

What prediction approach would you use?

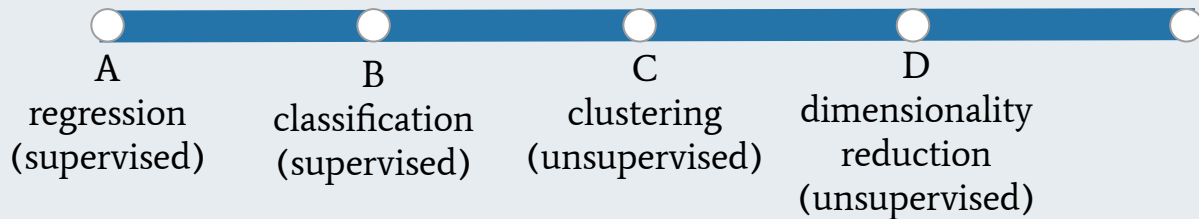


Prediction Approach



A real estate agency wants to predict the prices of houses based on features such as square footage, number of bedrooms, and location.

What prediction approach would you use?

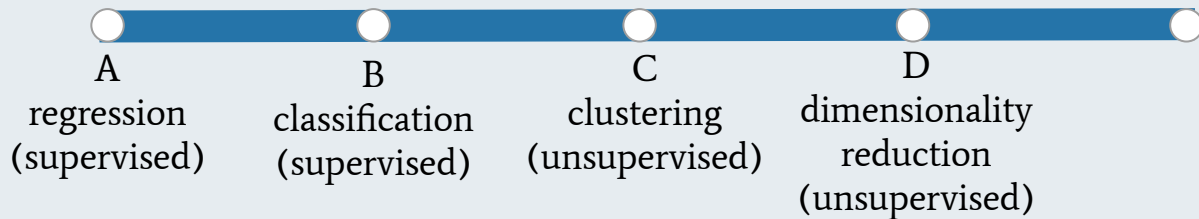


Prediction Approach

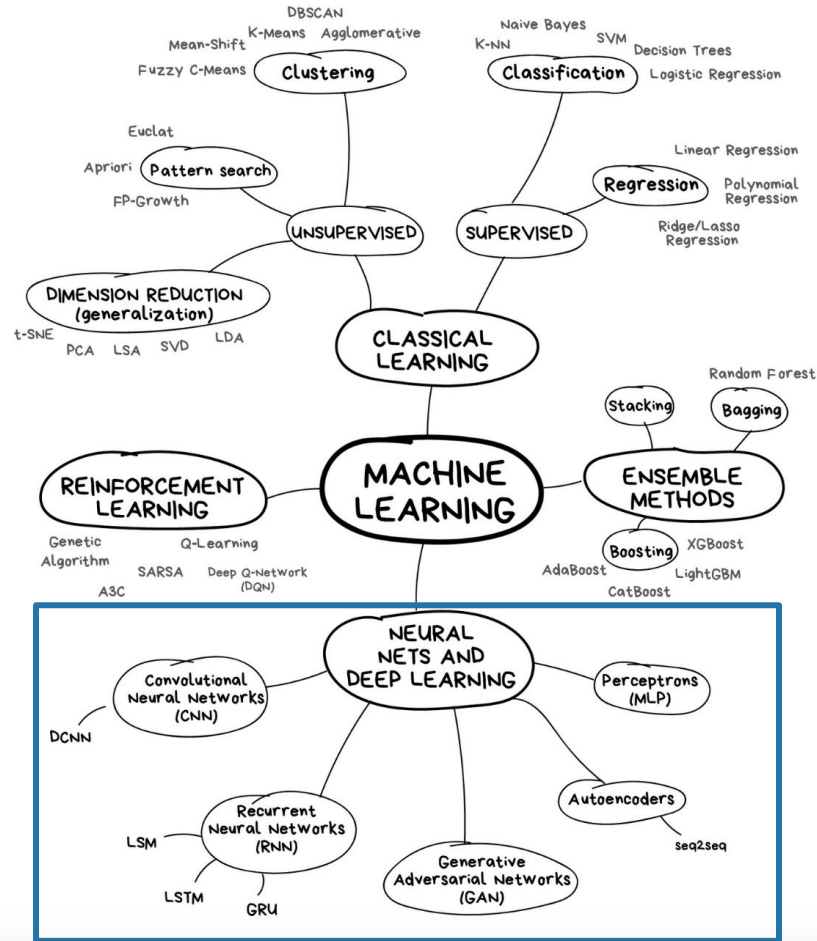


A retail store wants to organize its shelves by placing products with similar purchasing patterns together for better sales.

What prediction approach would you use?

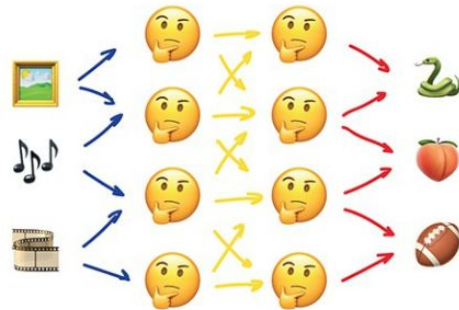


Deep Learning and Neural Networks

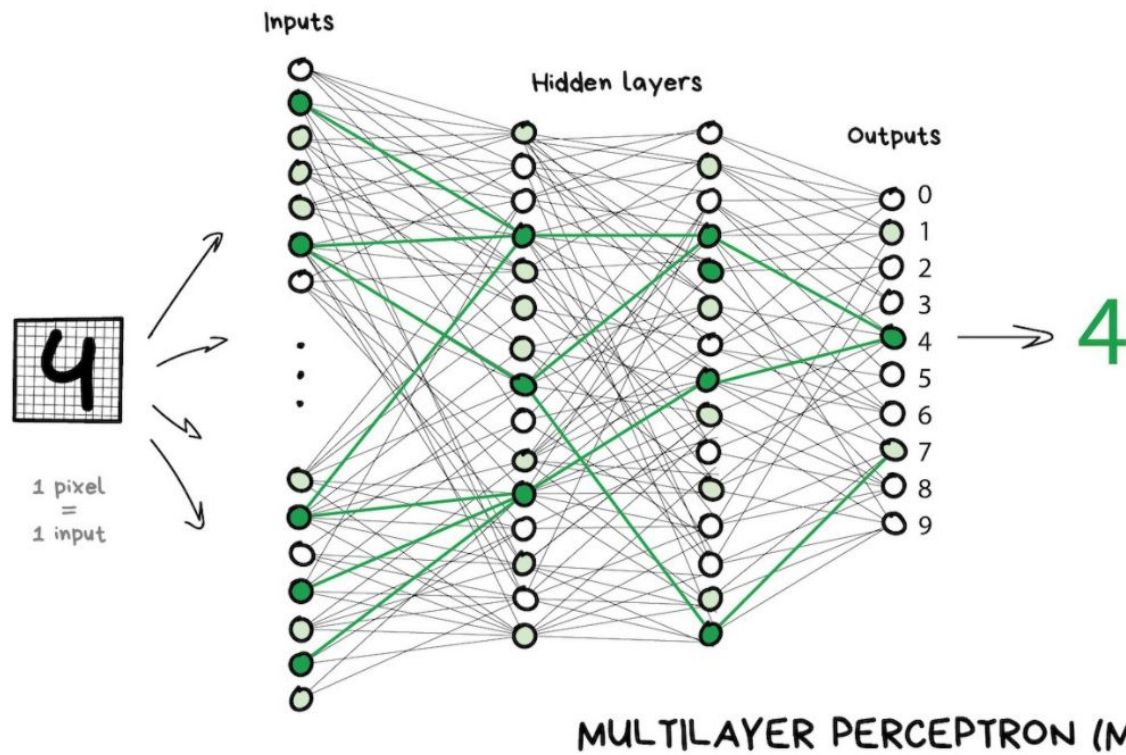


Neural Networks

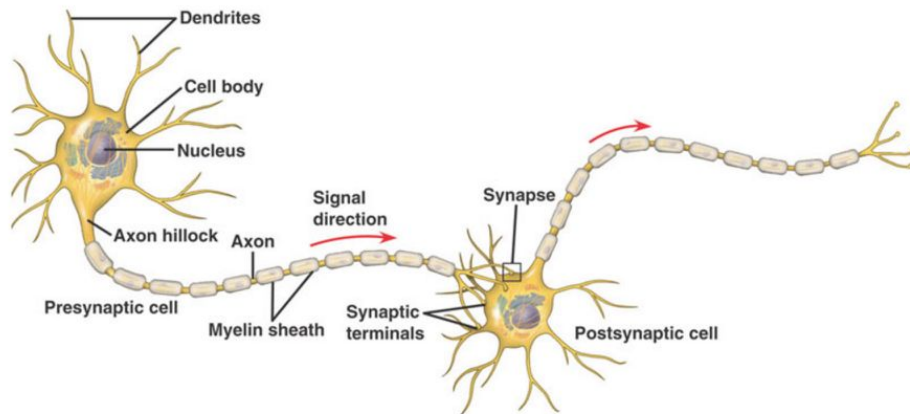
- Reduce high-dimensional data down to fewer dimensions
- Transforms data in a manner that preserves important, explanatory information
- Used for:
 - **Can** replace all the algorithms we've discussed
 - Image segmentation
 - Speech recognition
 - Image/text/video generation
- Popular neural network architectures include:
 - Multi-layer Perceptron
 - Convolutional Neural Networks
 - ResNet



Neural Networks

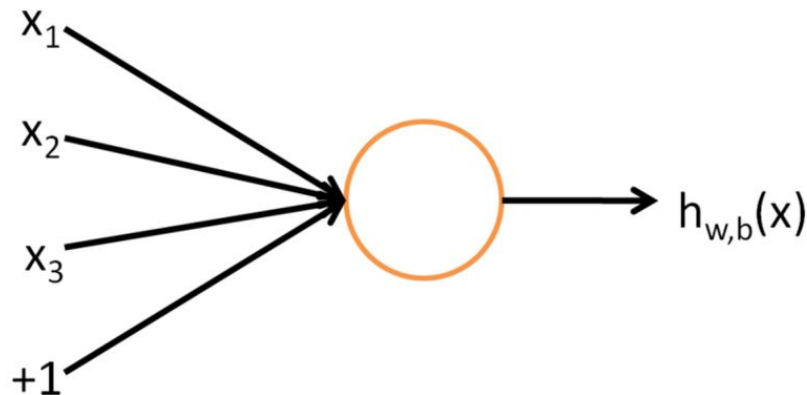


WHAT IS A NEURON?



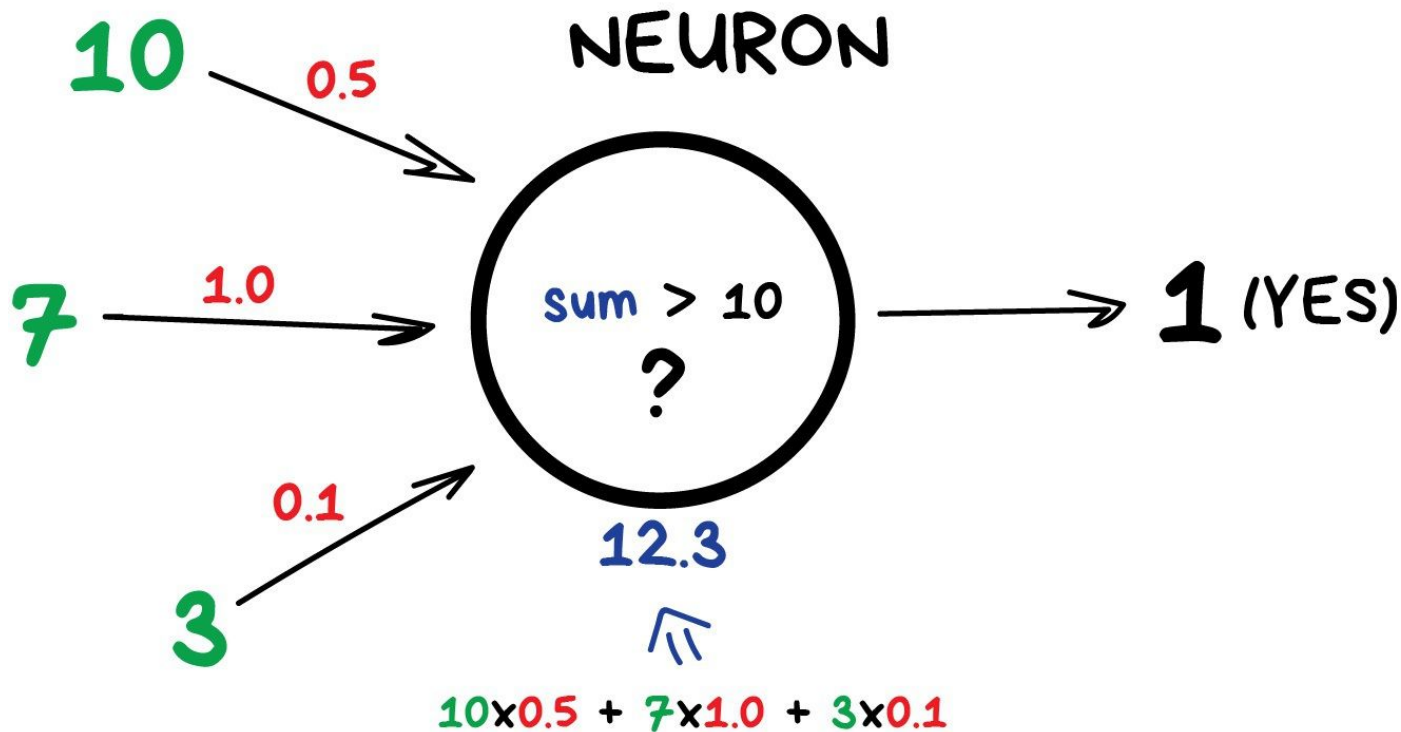
- Receives signal on synapse
- When trigger sends signal on axon

MATHEMATICAL NEURON



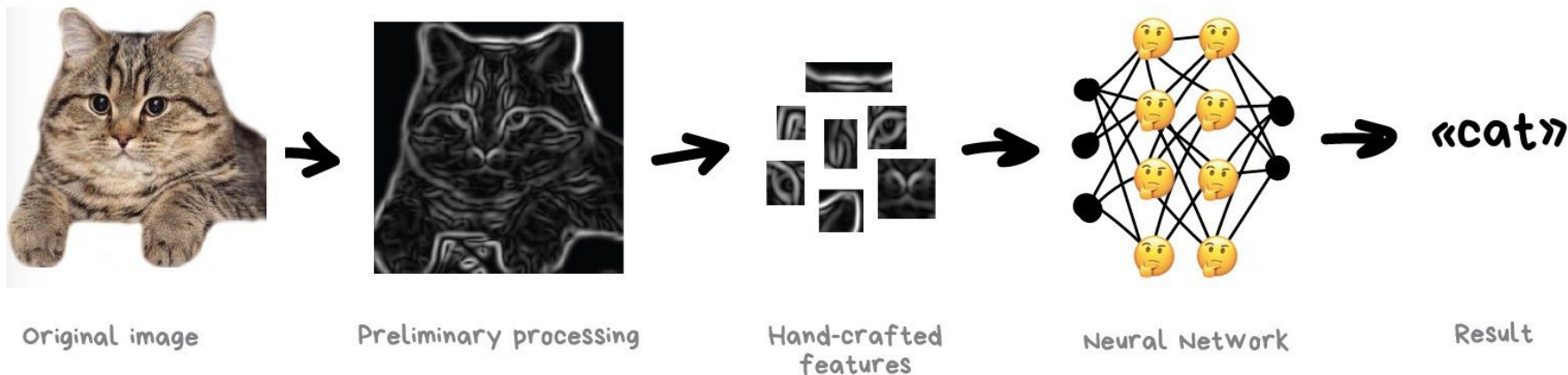
- Mathematical abstraction, inspired by biological neuron
- Either on or off based on sum of input

These weights tell the neuron to respond more to one input and less to another. Weights are adjusted when training — that's how the network learns. Basically, that's all there is to it.

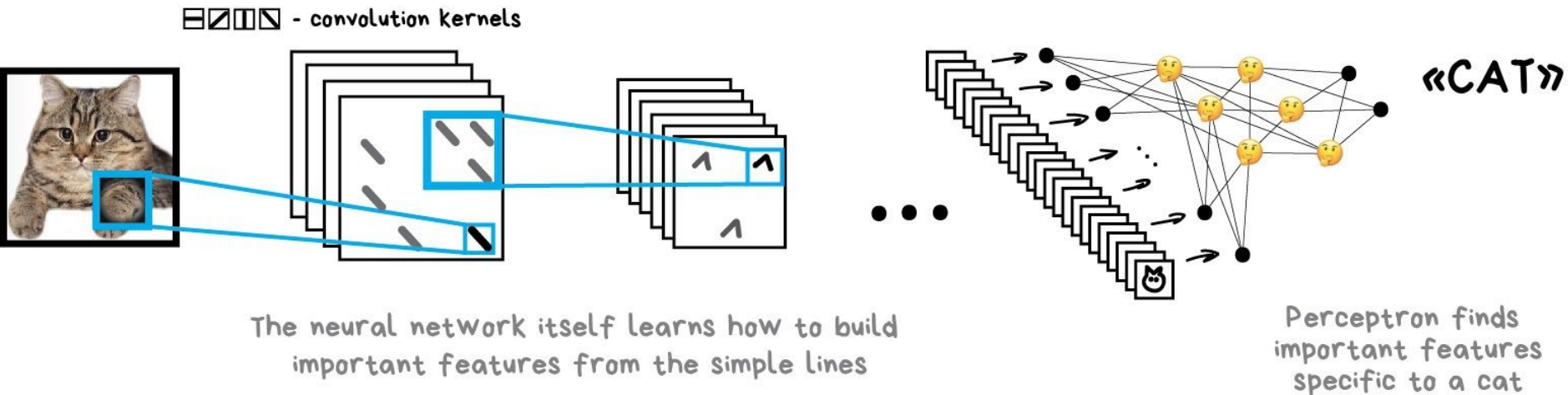


The challenges of image processing

- Extracting features from images is **hard**
 - Historically, humans would manually label images with particular features → handcrafting features



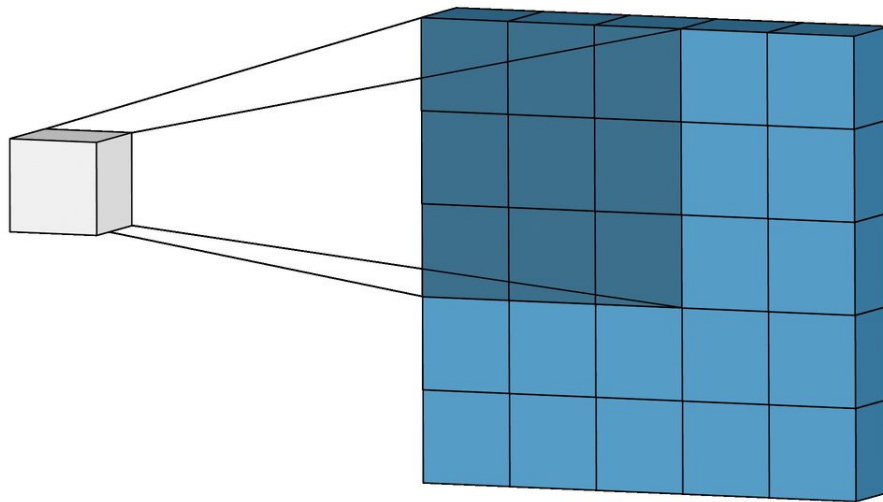
CNNs to the rescue!



“CNNs are all the rage right now. They are used to search for objects on photos and in videos, face recognition, style transfer, generating and enhancing images, creating effects like slow-mo and improving image quality. Nowadays CNNs are used in all the cases that involve pictures and videos.”

CONVOLUTIONAL NEURAL NETWORK (CNN)

CNNs to the rescue!



HOW A DEEP NEURAL NETWORK SEES

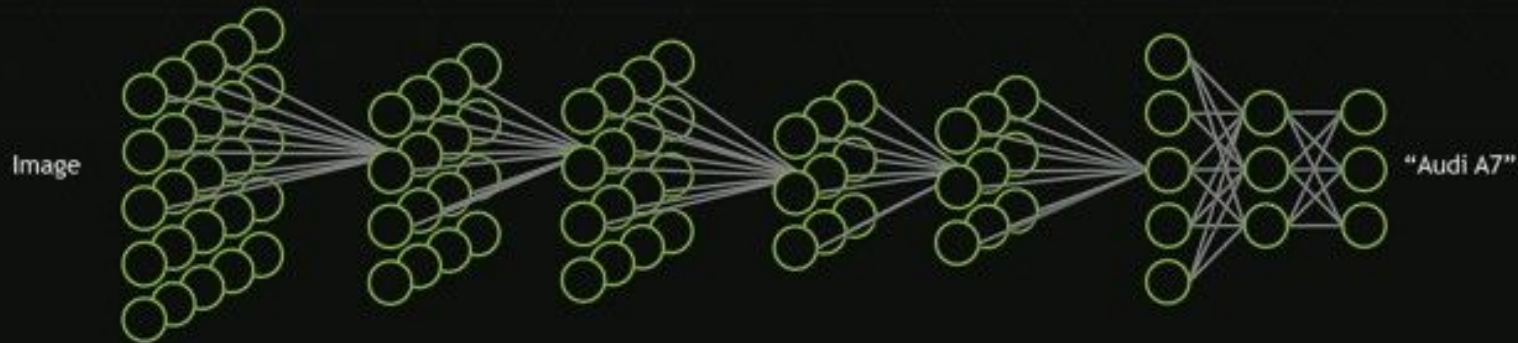
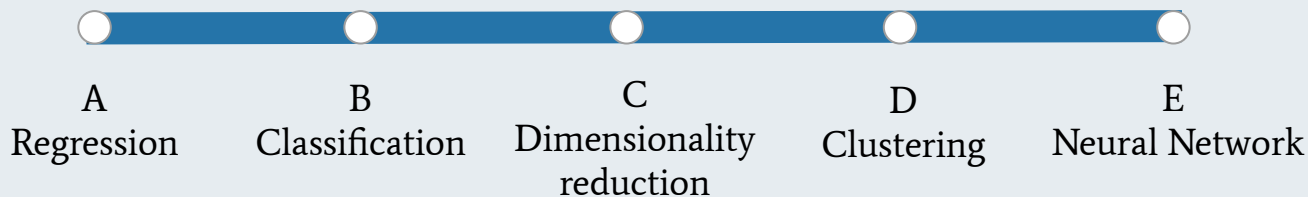


Image source: "Unsupervised Learning of Hierarchical Representations with Convolutional Deep Belief Networks" ICML 2009 & Comm. ACM 2011, Honglak Lee, Roger Grosse, Rajesh Ranganath, and Andrew Ng.

Reflection



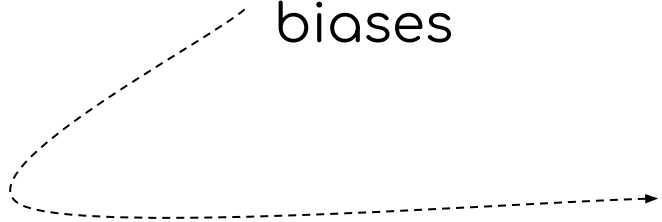
Which of the topics we covered was most interesting to you?



Beyond COGS 108

- **Supervised Learning - COGS 118A**
 - Labelled data, predict, classify, or fit a model
- **Unsupervised Learning - COGS 118B**
 - No labels, no feedback
 - Find hidden structures in a model
- **Reinforcement Learning - COGS 182**
 - No labels or rules; exploring an environment
- **Neural Networks - COGS 181**
 - Deep Learning and CNN architectures
- **Genetic Algorithms - COGS 186**
 - Mimicking evolution in an agent through trial-and-error

When models are trained
on historical data,
predictions will
perpetuate historical
biases



Predictive Analysis Ethics



Dare Obasanjo

@Carnage4Life

Product leader at Microsoft. My team is responsible for advertiser experience for Bing Ads; mobile apps, web UX, desktop apps & SDKs.



Dare Obasanjo

@Carnage4Life

Follow



Machine learning algorithms are driven more by the training data than math. Give an algorithm biased data then results will be biased. E.g.

- Amazon's resumé referral algo which auto rejected women
- Search ads algo which showed background check ads for "black sounding names"



Ryan Saavedra ✓ @RealSaavedra

Socialist Rep. Alexandria Ocasio-Cortez (D-NY) claims that algorithms, which are driven by math, are racist

8:59 PM - 22 Jan 2019

Amazon scraps secret AI recruiting tool that showed bias against women

Jeffrey Dastin

8 MIN READ



SAN FRANCISCO (Reuters) - Amazon.com Inc's ([AMZN.O](#)) machine-learning specialists uncovered a big problem: their new recruiting engine did not like women.

{* ARTIFICIAL INTELLIGENCE *}

MIT apologizes, permanently pulls offline huge dataset that taught AI systems to use racist, misogynistic slurs

Top uni takes action after *EI Reg* highlights concerns by academics

Katyanna Quach Wed 1 Jul 2020 // 10:55 UTC

SHARE









Janice Wyatt-Ross, Ed.D

@JaniceWyattRoss



Daughter 1 was taking an exam today being proctored by some type of software that apparently was not tested on dark skin. She had to open her window, turn on the lights, and then shine a flashlight over her head to be detectable. 😡😡😡

6:01 PM • Feb 22, 2021 • Twitter for iPhone

7,030 Retweets **939** Quote Tweets **34.6K** Likes

What to do about bias...

1. Anticipate and plan for potential biases before model generation. Check for bias after.
2. Have diverse teams.
3. Use machine learning to improve lives rather than for punitive purposes.
4. Revisit your models. Update your algorithms.
5. You are responsible for the models you put out into the world, unintended consequences and all.

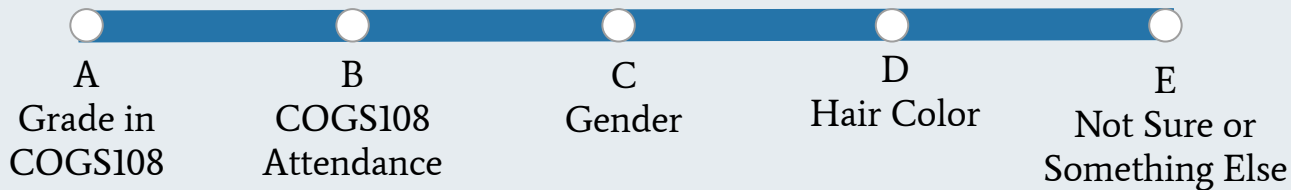
Data Science Question

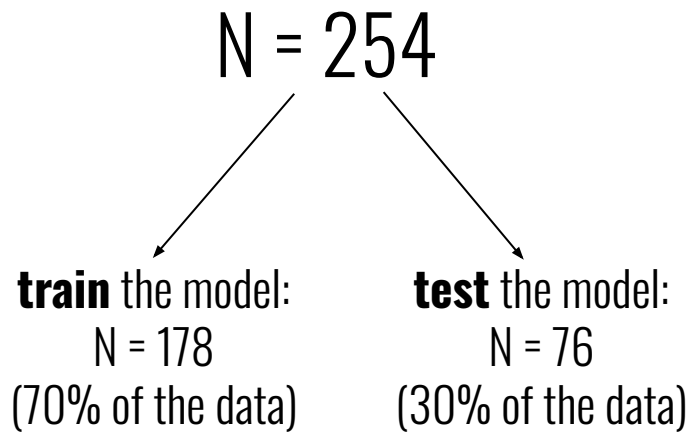
Based on data I have about you all, can I predict
who in this course will be successful?

Prediction Approach



Which would be the most predictive of your future success?



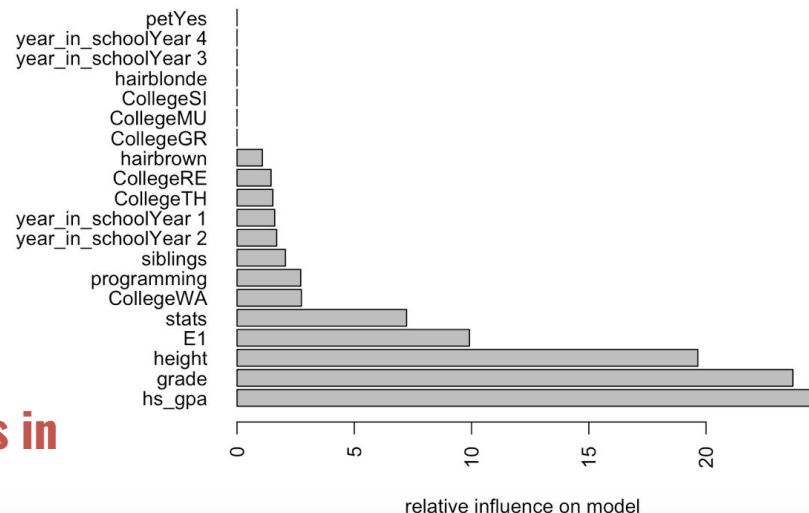


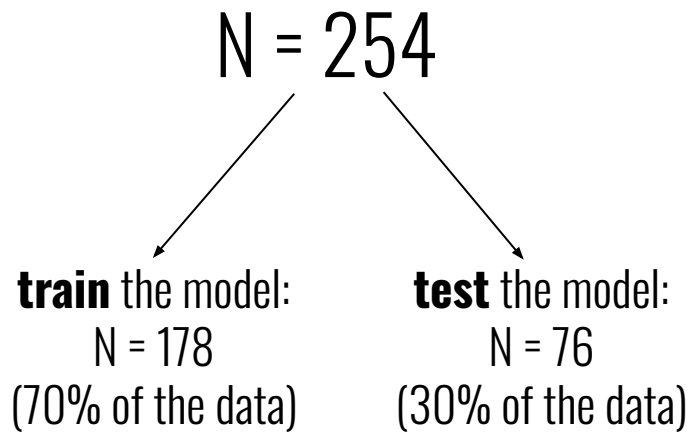
train the model

**predicted success in
test set**

	Accuracy	Sensitivity	Specificity
training set	71.2%	76%	67%
test set	49.1%	40%	60%

Assess Prediction Model



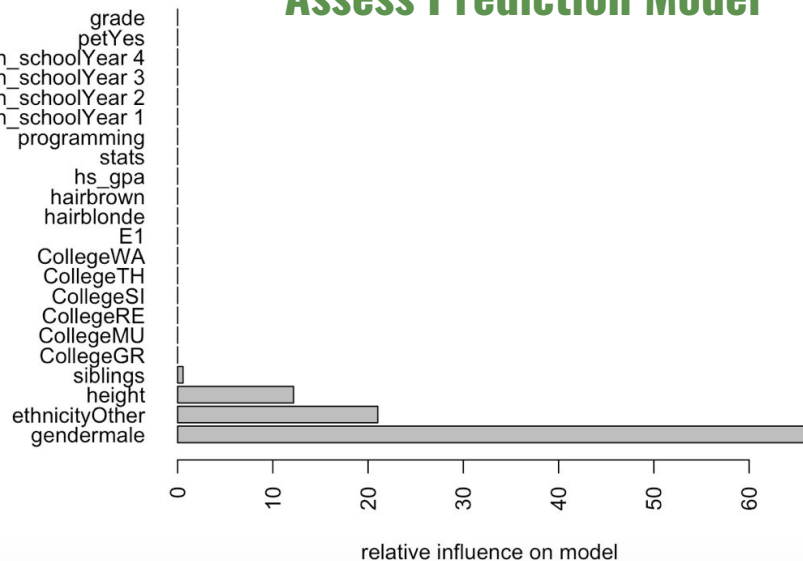


train the model

**predicted success in
test set**

	Accuracy	Sensitivity	Specificity
training set	100%	100%	100%
test set	100%	100%	100%

Assess Prediction Model



What if I were using these data to determine who I should write recommendation letters for?

Or to determine which students I focus my attention on?

Or whose projects I read?

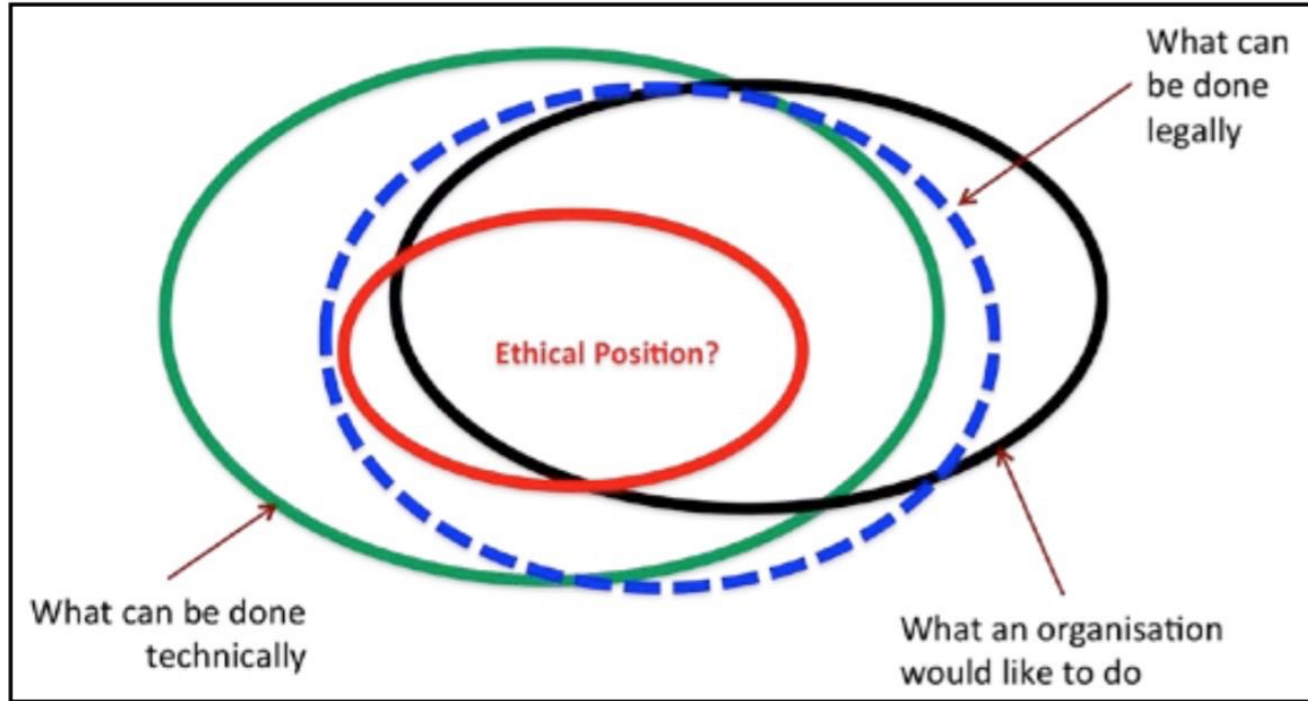
Or who I allow to come to office hours?

Or who UCSD allows to be data science majors?



Think about whether the models you're building should even be built.

Big Data Ethics



Predictive algorithms should (*at a minimum*) be FAT

Fair: lacking biases which create unfair and discriminatory outcomes

- For whom does this algorithm fail?
- Steps to take:
 1. Verify data about individual is correct
 2. Carry out “sensitivity test”

Accountable/Accurate: answerable to the people subject to them

- Correct data used? Is there a mechanism for appeal?

Transparent: open about how and why particular decisions were made

- Think *carefully* about what transparency is (Handing over source code likely isn't the answer)

A Mulching Proposal

Analysing and Improving an Algorithmic System for Turning the Elderly into High-Nutrient Slurry

Os Keyes

Department of Human Centered Design & Engineering
University of Washington
Seattle, WA, USA
okeyes@uw.edu

Meredith Durbin

Department of Astronomy
University of Washington
Seattle, WA, USA
mdurbin@uw.edu

Jevan Hutson

School of Law
University of Washington
Seattle, WA, USA
jevanh@uw.edu

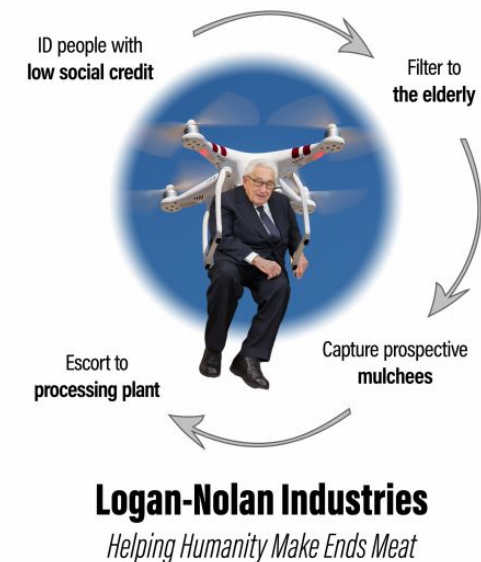


Figure 1: A publicity image for the project, produced by Logan-Nolan Industries

A Mulching Proposal - Fair

Problem

- Reference datasets disproportionately tagged white cisgender men as worthy of mulching, biasing against other populations

Solution

- Collected photographs and social credit data of 3,000 more potential mulchees, particularly women, trans people and/or people of color

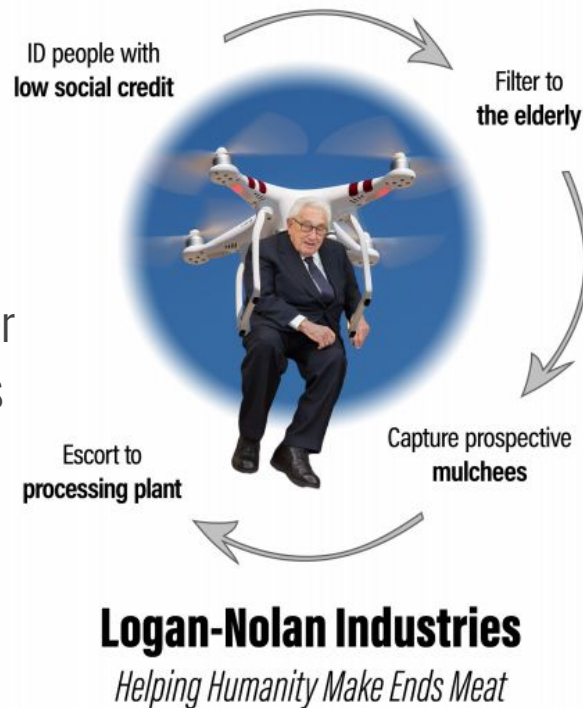


Figure 1: A publicity image for the project, produced by Logan-Nolan Industries

A Mulching Proposal - Accurate

Problem

- A computer vision algorithm could fail
- Analysis of social connections might be inaccurate

Solution

- Soliciting feedback from mulchees and their relatives/friends at various stages of the mulching process
- Drones gives window of time for mulchee to appeal their case via a phone call
- If incorrectly mulched, they will be compensated

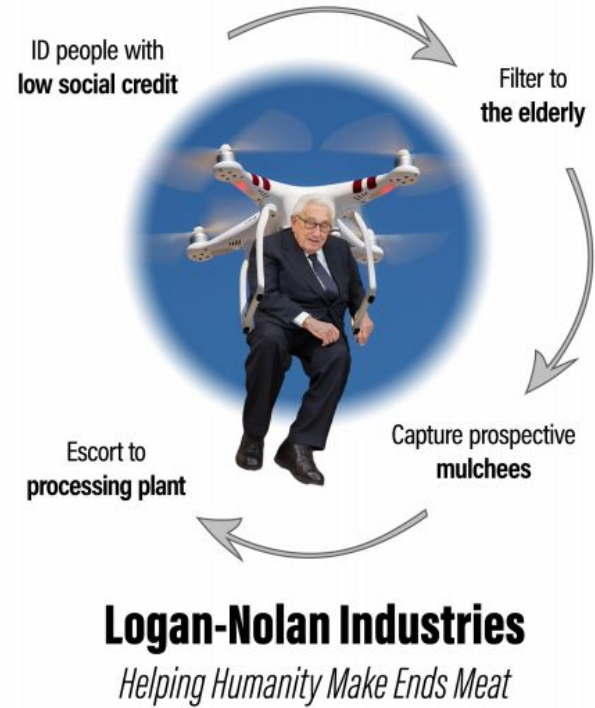


Figure 1: A publicity image for the project, produced by Logan-Nolan Industries

A Mulching Proposal - Transparent

Problem

- People should be aware of the decisions and procedures being conducted behind the system

Solution

- All variables used will be made public: including by not limited to phone and SMS metadata, number of birthday cards received from relatives—along with mulchee scores
- Allow people to play with settings and see who would be mulched

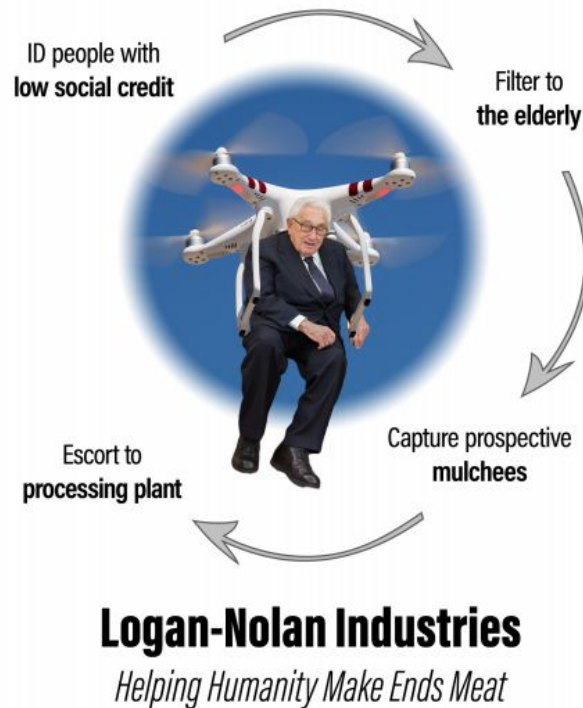


Figure 1: A publicity image for the project, produced by Logan-Nolan Industries

A Mulching Proposal

FAIR - equally considers all elderly individuals

ACCURATE - pre- has mechanism for appeal; post - compensation

TRANSPARENT - website with all features; testable

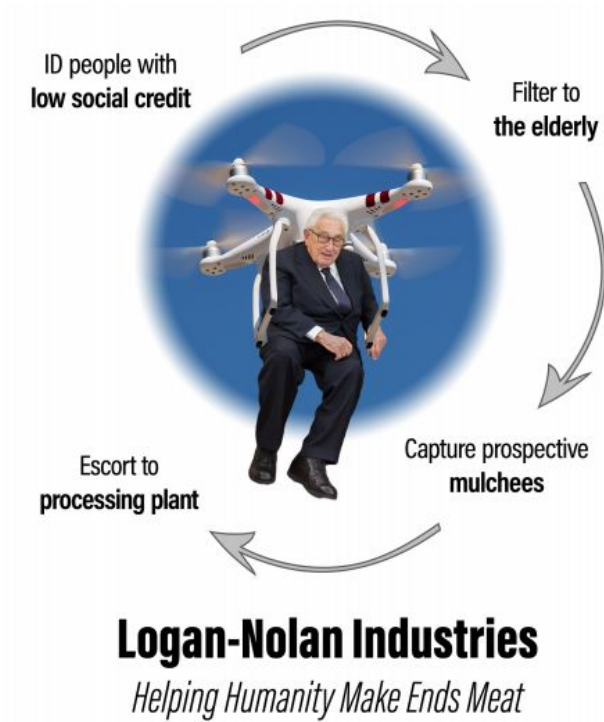
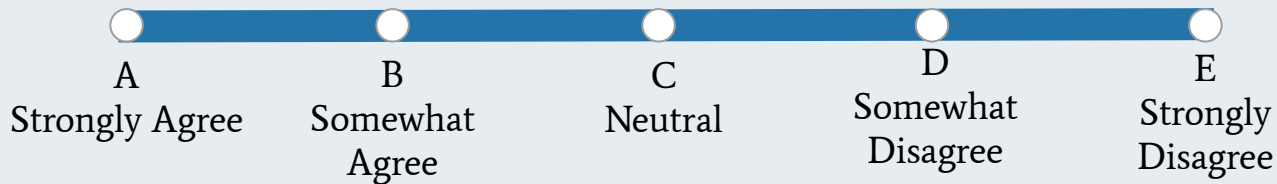


Figure 1: A publicity image for the project, produced by Logan-Nolan Industries

Prediction Thoughts



We should start using this algorithm to mulch up the elderly



Checklists are helpful, but they're not an excuse for thoughtlessness.

Thank you
