LESSON 5: SPECIAL TOPICS AND DISCUSSION

Savannah Thais

Intro to Machine Learning

Princeton Wintersession 2021

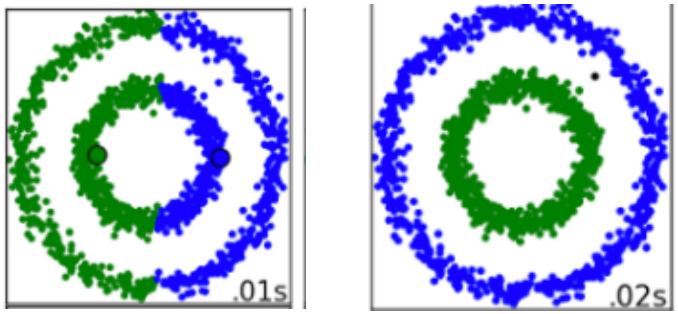
01/29/2021

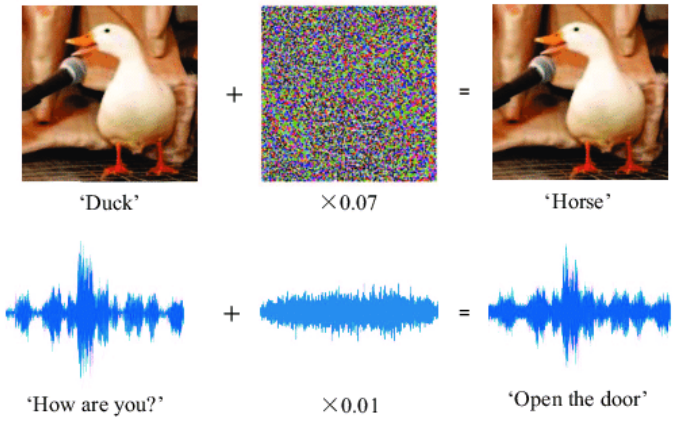
Outline for Today

- Quick Review
- A Scientific Approach to ML
- Research Applications Discussion
- Intro to AI Ethics/Fairness
- Open Discussion

Knowledge Review

- What are the two components of a GAN?
- Which type of clustering were probably used in these examples?
- What is the goal of an autoencoder?
- What is happening in this diagram?
- What are Deep Fakes?
- Why is dimensionality reduction useful?



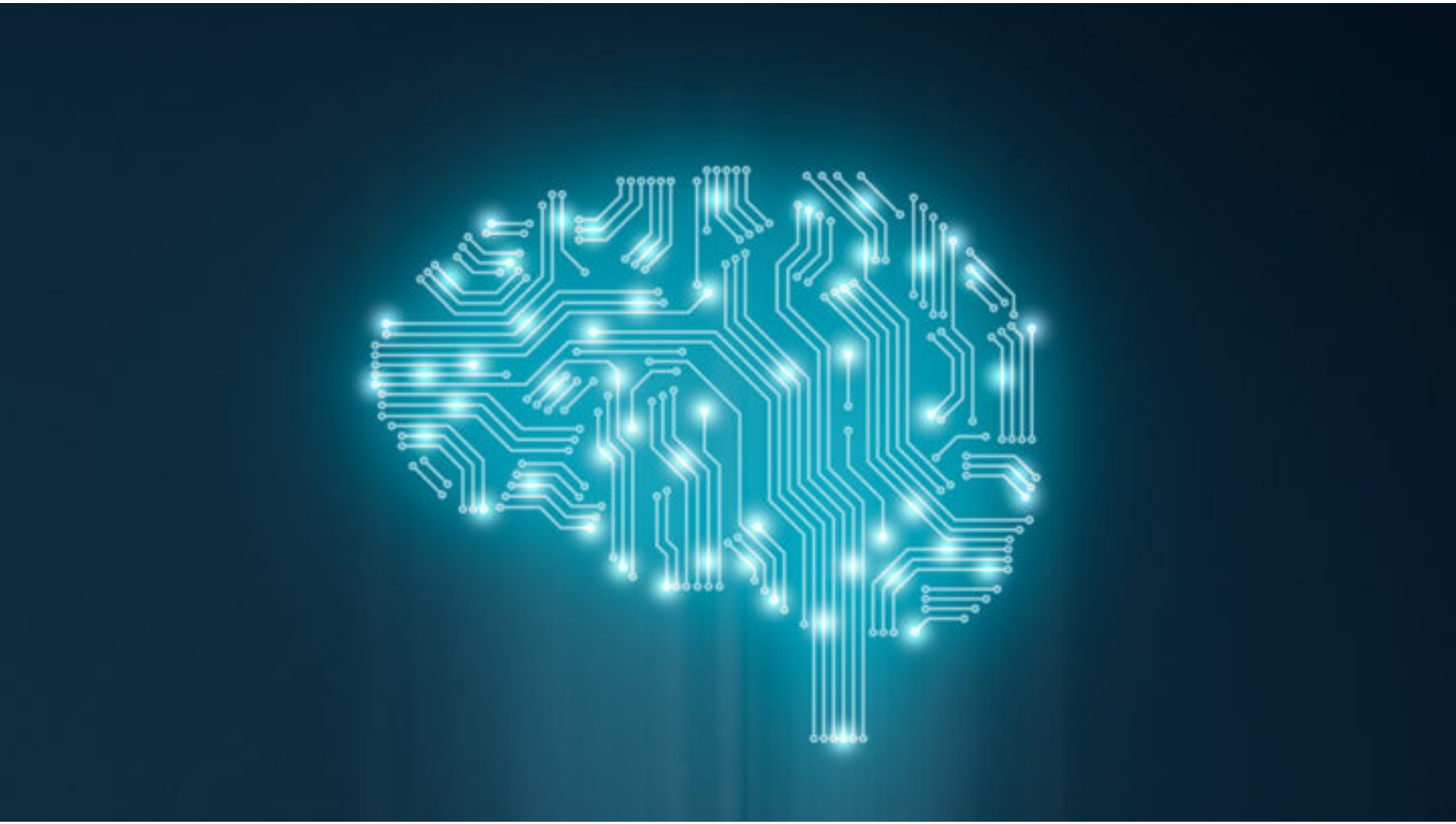


Let's revisit a question from Monday

Can you think of an example of ML in your daily life (or research)?:

- Do you think it's a supervised or un-supervised model?
  - What type of model do you think was used?
  - What data could have been used to train it?
  - What might have been the learning objective?
- Can you imagine some important training considerations?

A Scientific Approach to ML



Conceptual Review

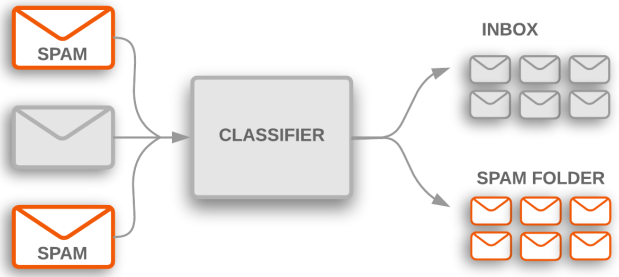
What are some examples of problems ML are well suited for?

- What are some problems ML is NOT suited for?
- What are some important considerations during model development and training?

What are Problems for ML?

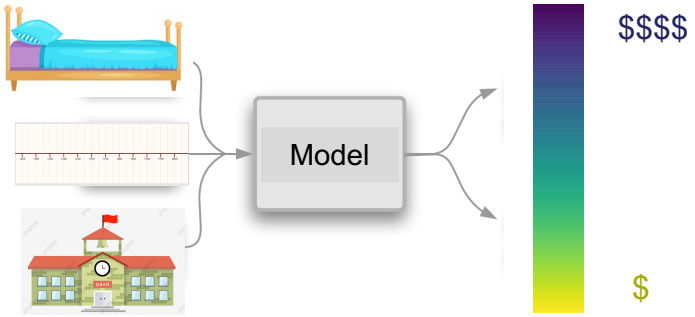
Classification:

Predict a class label for an input



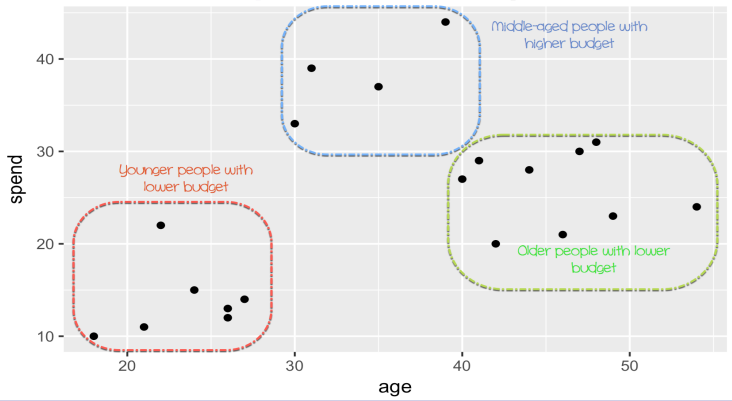
Regression:

Predict a continuous variable



Clustering:

Group similar inputs



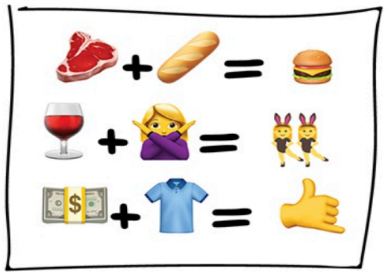
Generation:

Construct new data within pattern

What are Problems for ML?

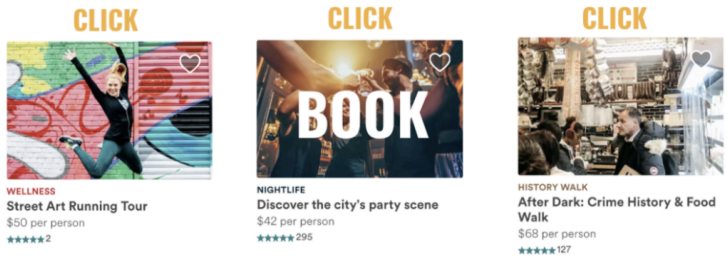
Association Rules:

Identify common patterns in data



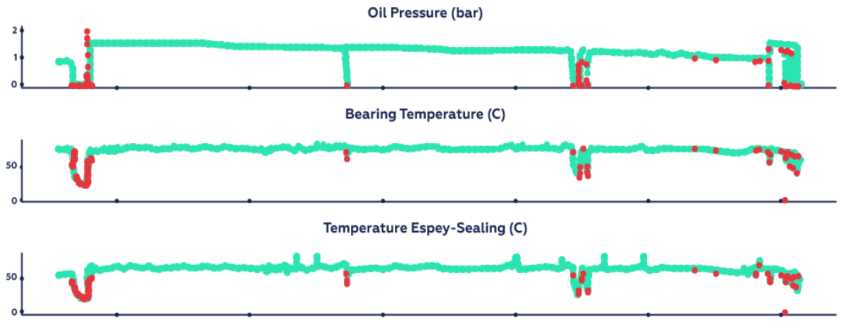
Ranking:

Generate optimal orderings



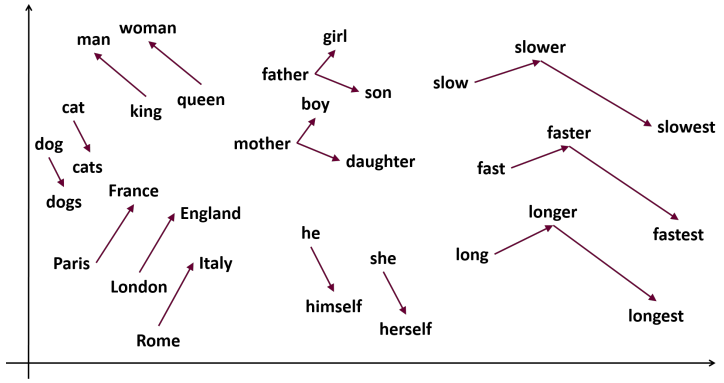
Anomaly Detection:

Identify statistical outliers



Restructuring:

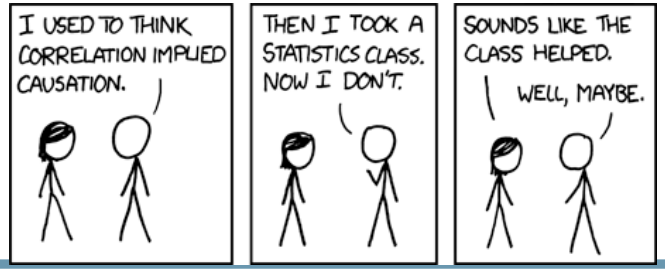
Transform data representations



What are NOT Problems for ML?*

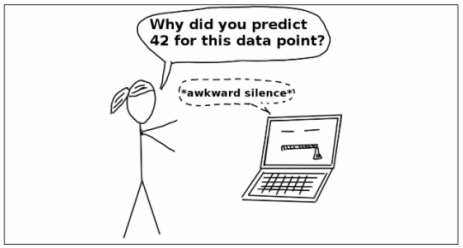
Causation:

Models learn correlations, but can't infer causality or intent



Precise Interpretability:

It's often difficult to understand what a model is learning



Context:

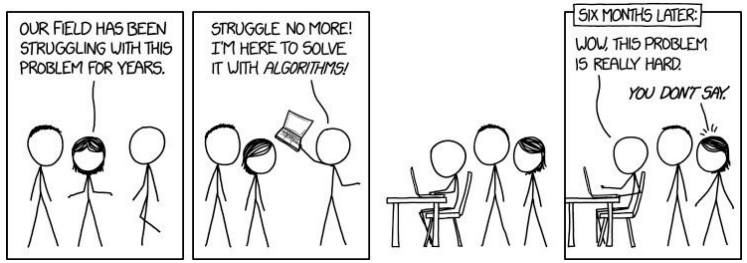
Models are incapable of non-mathematical reasoning

Keaton Patti ✓
@KeatonPatti

I forced a bot to watch over 1,000 episodes of Jerry Springer and then asked it to write an episode of its own. Here is the first page.

Data Limitations:

Models can't fix problems in data or learn without examples



ML Should Be Scientific!

Designing a (good) ML model is like running a scientific experiment: we don't know apriori what will work best

Step	Example
1. Set the research goal.	I want to predict how heavy traffic will be on a given day.
2. Make a hypothesis.	I think the weather forecast is an informative signal.
3. Collect the data.	Collect historical traffic data and weather on each day.
4. Test your hypothesis.	Train a model using this data.
5. Analyze your results.	Is this model better than existing systems? *
6. Reach a conclusion.	I should (not) use this model to make predictions, because of X, Y, and Z.
7. Refine hypothesis and repeat.	Time of year could be a helpful signal.

* Including how certain you are!

source

The Hypothesis

Your ML hypothesis is a combination of the model you want to build and the pattern you want to explore

- "An algorithm can distinguish between normal and cancerous brain scans based only on pixel values"
- "A model can simulate tau lepton decays within a defined margin of uncertainty"

- Questions to consider as you construct your hypothesis:
  - What specifically do I want my model to be able to do?
  - What is the ideal outcome/use case of my experiment?
  - What will I consider a success (proving hypothesis) or failure (rejecting hypothesis)?
  - What kinds out outputs do I need the model to make and how will I use them?

The Experiment

Building, training, and evaluating your model is the experimental process of testing your hypothesis

- Your learning goal, input data, and desired output structure can help determine what class of models to study
- All components need to be quantifiable and measurable
  - What are your input features and how are they represented?
  - What is the specific learning task for the model?
  - How do you quantify how well the model is doing?
  - What metric can you use to compare different models?

Setting Up Your Data

Your model is only as good as your training data

How much data is available and does each entry have the same information?

Do you have examples of all data classes/ranges?

How expensive is it to create/collect more data or labels?

Are the available labels related to the decision you want to make?

Is there noise in your label creation or distribution?

Are classes and inputs balanced and normalized?

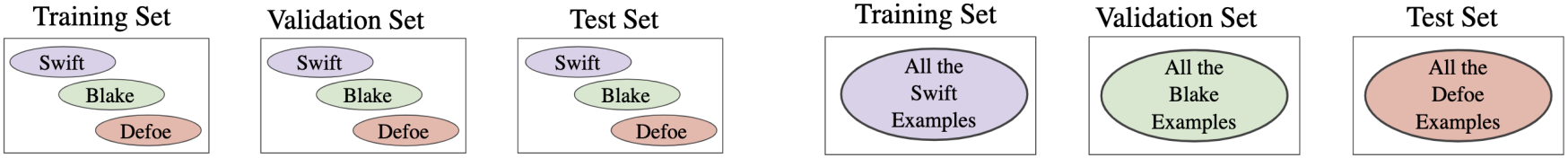
Are there patterns in your data you don't want the model to exploit?

ML and Research

Example Problems

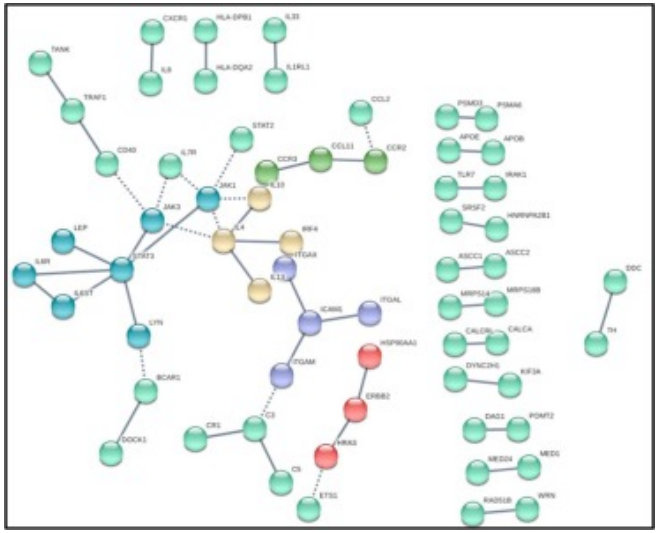
Predicting author's political associations from 'mind metaphors'

- Supervised multi-class classifier
- Each metaphor (embedded) is an input data point
- Evaluate based on accuracy of predictions (% of correct classifications)
- Watch out! Could learn to associate writing styles with individual authors



Distinguishing genetic cohorts

- Unsupervised clustering
- Patient's full genome as input data
- Evaluate based on measurable differences (disease manifestation) between clusters and gene pathway analysis of differences



Computer Vision for Disaster Response

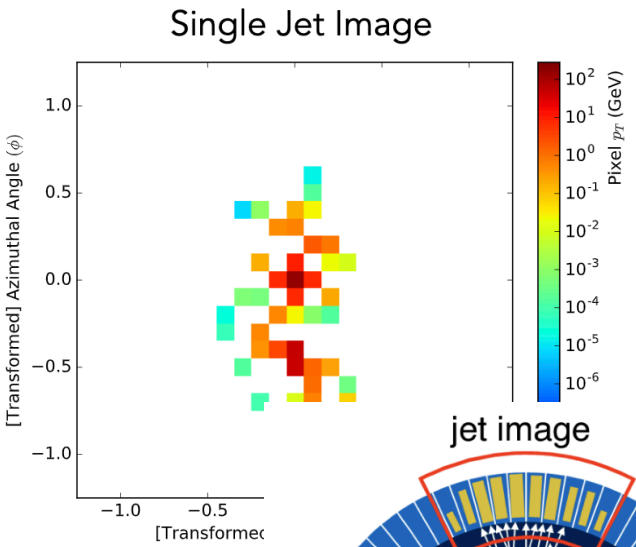
Use a multi-model ML pipeline to localize damage after a natural disaster and prioritize response

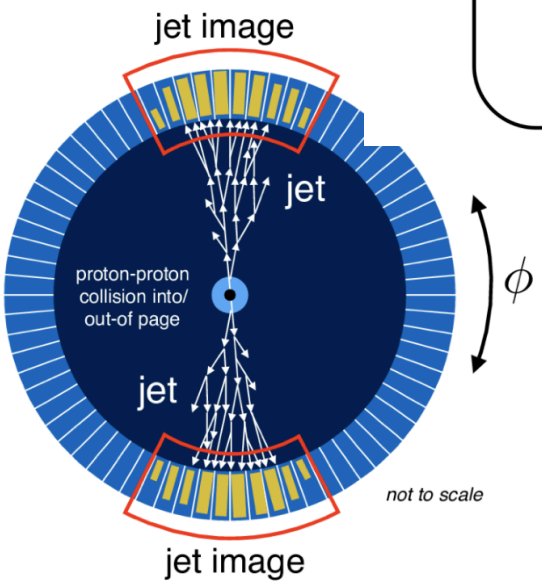
[Blog Post]

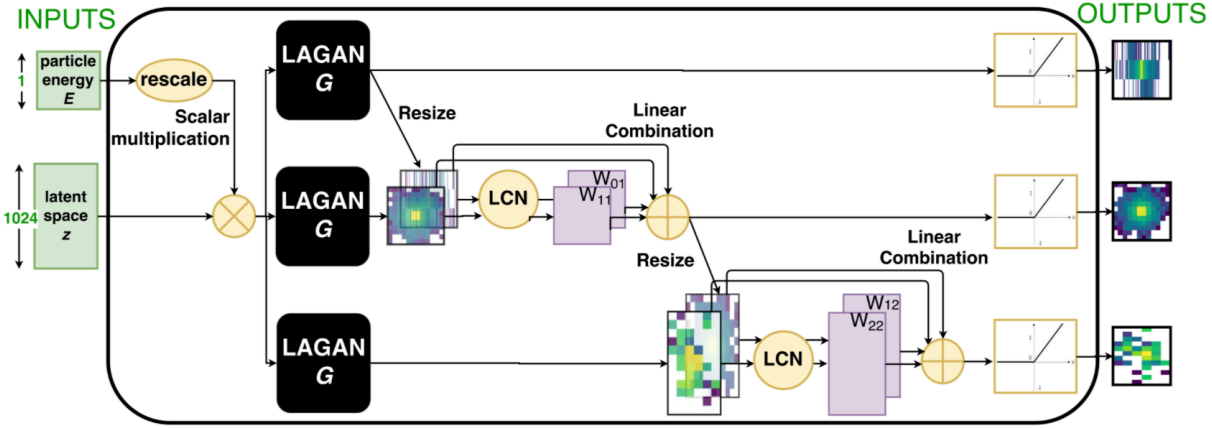
[Academic Paper]

GANs for Particles

Use a generative model to accelerate particle physics simulation







[GitHub]

[Talk]

[Academic Paper]

Your Turn!

Think of a research problem where you want to employ ML:

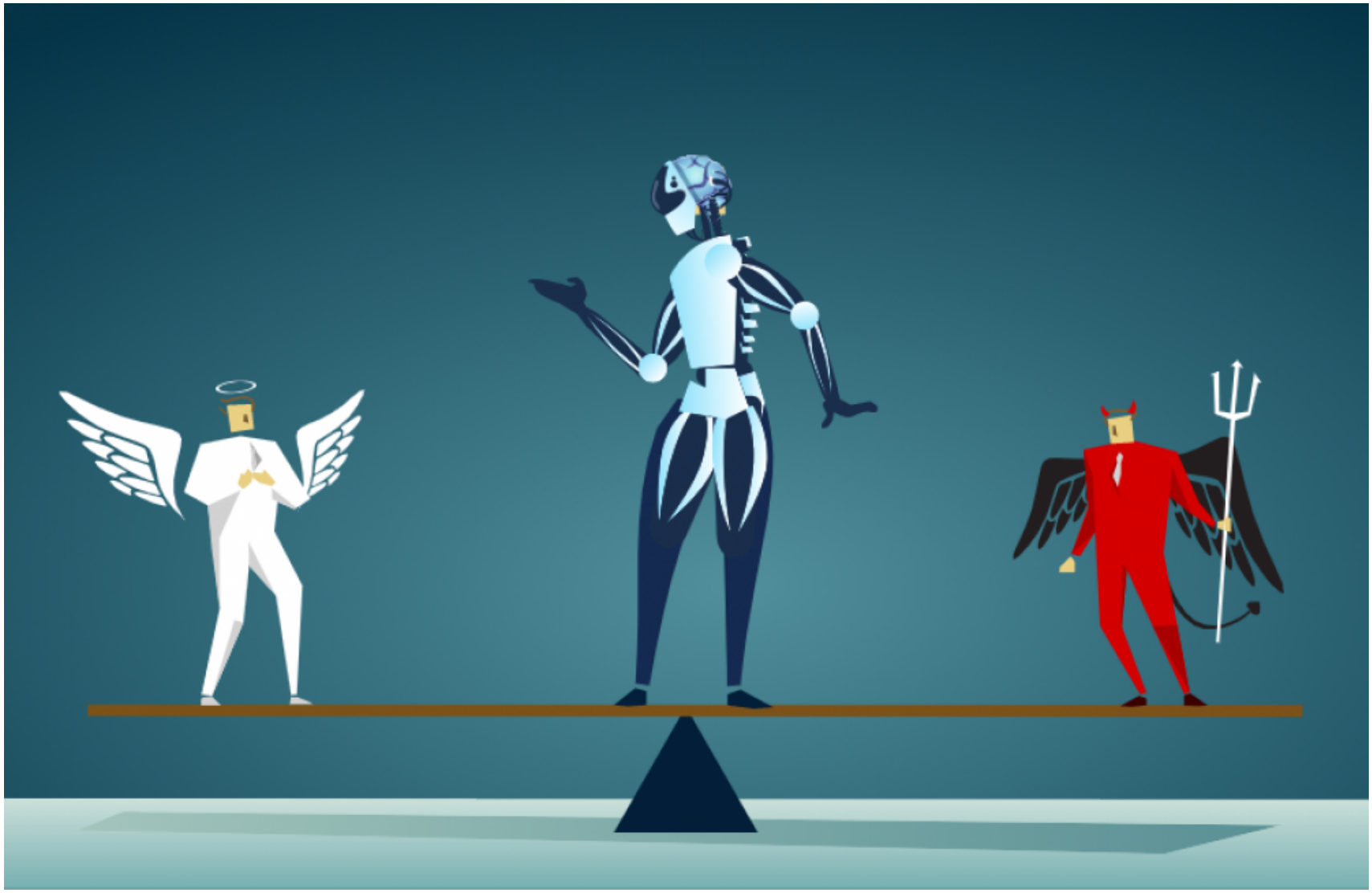
- What is your (testable!) hypothesis?
- What data could you use as inputs?
- What would the learning goal be?
- How would you quantify success?

What model(s) would you use in your experiment?

- What parameters would you need to consider?
- Would you have enough/the right training data?
- Would the model enable useful research decisions?

Break

A Discussion Primer on AI Ethics



Data Biases

- Apple Pay card gave higher (or any) credit limits to men
  - Models trained on historical data may be ‘accurate’ but not ‘fair’
- Kidney function measurement algorithm (eGFR) misdiagnoses African American patients
  - Misunderstanding reasons for group differences in training data can have drastic results
- COMPAS recidivism prediction tool predicts higher scores for minorities
  - Messy correlation between who actually recidivates and who is targeted by law enforcement
  - Impossible to make fair?

Unequitable Application

- Using facial recognition entry in rent-stabilized housing
  - Facial recognition software is often tuned for white faces, using it in low income communities can be an effort towards gentrification
- Rite Aid deployed facial recognition in low income areas
  - Systems are deployed on people they're not built for, who don't have a say in their development, and don't opt in

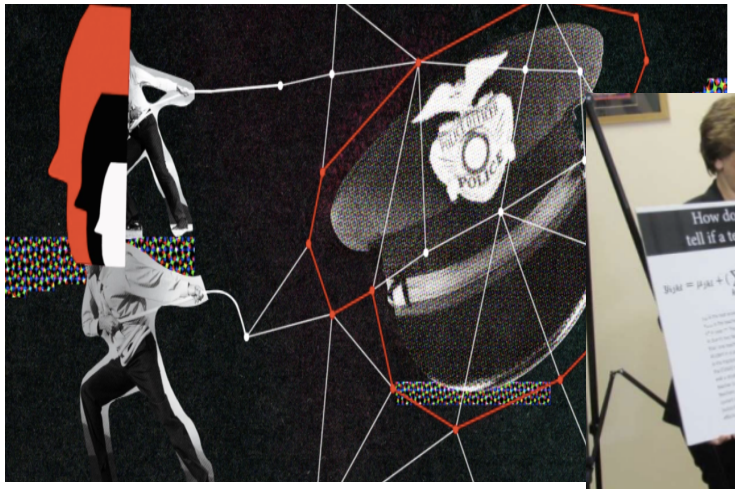
In the hearts of New York and metro Los Angeles, Rite Aid installed facial recognition technology in largely lower-income, non-white neighborhoods, Reuters found. Among the technology the U.S. retailer used: a state-of-the-art system from a company with links to China and its authoritarian government.

BIG CITY

The Landlord Wants Facial Recognition in Its Rent-Stabilized Buildings. Why?

Bias, Targeting, Regulation

- Chinese government using facial recognition and race classification algorithms to target Muslims
- Predictive policing algorithms target neighborhoods with higher police activity, regularly mis-classify people
- Government utilize Automated Decision Systems are not auditable or tested for accuracy and bias



A Note on Algorithmic Fairness

What is an Algorithmic Practice Audit?

An independent, third party review of an organization's algorithmic processes and outcomes

SCOPE

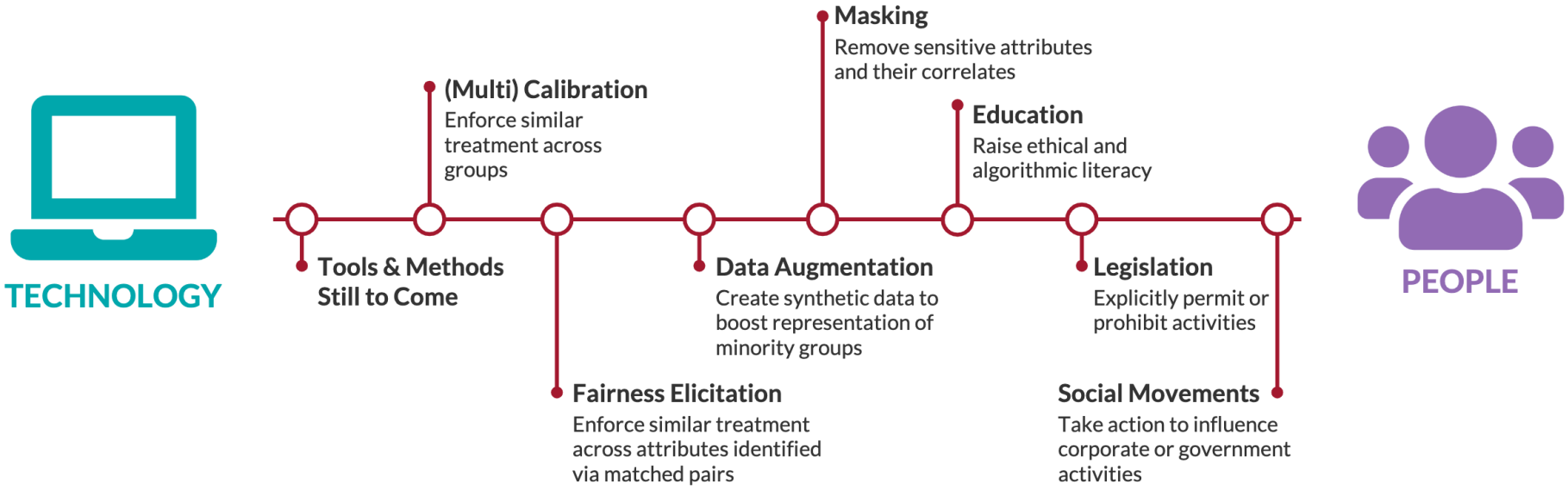
- Process
  - Is training data representative?
  - Does data cleaning / presentation introduce bias?
  - Are fair classes of algorithms used?
- Outcomes
  - Does the model meet its stated fairness goals?
  - Is there disparate impact or measurable bias?
  - Is bias introduced by humans in the "last mile"?

BENEFITS

- Signal to consumers and (shareholders) that algorithmic services are correct and fair
- Use a forcing function to improve internal processes and controls
- Take pride in certification that you're doing the right thing

Excellent full talk!

A Multi-Faceted Approach



Excellent full talk!

Open Discussion

Happy to answer any questions or explore additional topics!

sthais@princeton.edu @basicsciencesav