

Machine Learning for Data Science (CS4786)

Lecture 27

Socially Responsible Machine Learning

Course Webpage :

<http://www.cs.cornell.edu/Courses/cs4786/2017fa/>

Announcements

- Survey 2, over 90%.
- Kaggle date is not changeable.

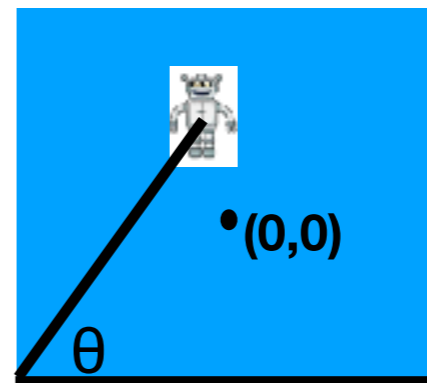
Some Tips for Competition 2

Some Tips for Competition 2

- Before you try continuous state models, first try the finite state one

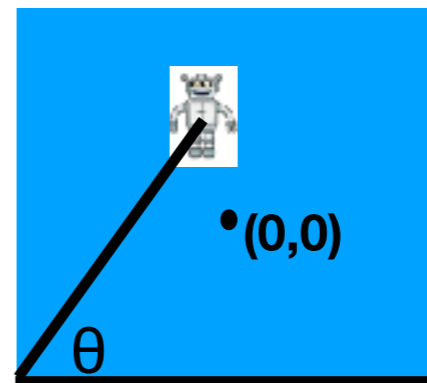
Some Tips for Competition 2

- Before you try continuous state models, first try the finite state one
- Angle to bottom edge



Some Tips for Competition 2

- Before you try continuous state models, first try the finite state one
- Angle to bottom edge



- If you use angles as observations you can learn HMM model

Some Tips for Competition 2

Some Tips for Competition 2

- To get an estimate of how many states you might need, analyze given sample locations

Some Tips for Competition 2

- To get an estimate of how many states you might need, analyze given sample locations
- Learn to map states to locations based on data given to you and use this to make predictions

Some Tips for Competition 2

- To get an estimate of how many states you might need, analyze given sample locations
- Learn to map states to locations based on data given to you and use this to make predictions
- Other thing to try (only after you get basic model to work): What if we are given state variables for some intermediate time steps

Some Tips for Competition 2

- To get an estimate of how many states you might need, analyze given sample locations
- Learn to map states to locations based on data given to you and use this to make predictions
- Other thing to try (only after you get basic model to work): What if we are given state variables for some intermediate time steps
 - what happens to forward-backward steps, first work it out on pen and paper

Some Tips for Competition 2

- To get an estimate of how many states you might need, analyze given sample locations
- Learn to map states to locations based on data given to you and use this to make predictions
- Other thing to try (only after you get basic model to work): What if we are given state variables for some intermediate time steps
 - what happens to forward-backward steps, first work it out on pen and paper
 - hint: you might not even need to reimplement F pass

ML DREAM

shoes

All Shopping Maps Images News More Settings Tools

About 1,810,000,000 results (1.10 seconds)

Shoes at Zappos.com
Ad www.zappos.com/Shoes
4.7 ★★★★★ rating for zappos.com
Fast, Free Shipping & Free 365 Day Returns on Huge Selection of Shoes!
Birkenstock · Nike · Converse · New Balance · Born · Frye
Types: Sneakers, Slippers, Heels, Boots, Flats, Running Shoes


Womens Shoes at Macy's - Save 40-60% on Cyber Monday - macys.com
Ad www.macys.com/Womens_Shoes/Holiday_Deals
4.2 ★★★★★ rating for macys.com
Cyber Monday Ends 11/30, Deals Going Fast + Extra 20% Off!
Styles: Boots, Wedges, Loafers, Flats, Slippers, Sneakers, Pumps, Espadrilles, Booties
Ratings: Returns 9/10 - Product quality 9/10 - Shipping 8.5/10 - Service 8.5/10 - Selection 8.5/10

Converse® Official Site - Converse.com
Ad www.converse.com/Shoes
Full Converse Collection. Shop Our New Designs for Men, Women & Kids.
Free Shipping for Members · Free 60 Day Returns
Types: Men's Sneakers, Women's Sneakers, Kid's Sneakers, Custom Sneakers

Brand Name Shoes for Less - Up to 70% Off - zulily.com
Ad www.zulily.com/
Exclusive Deals On All Shoes, Socks, & More at Zulily. Shop & Save Today!

Men and Womens shoes, Shipped Free | Zappos.com
www.zappos.com/shoes
Boots Sneakers & Athletic Heels Flats Sandals View All... Sneakers & Athletic Boots Oxfords Loafers Sandals View All... Sneakers and Athletic Boots Slippers Flats Sandals view all...
Sandals · Women's Shoes · Sneakers & Athletic Shoes · Popular Men's Shoe Styles

Shop on Google Sponsored ⓘ

 Adidas NEO Baseline Women's Dicast-Leather Sneakers, Size: 8, White
\$34.99 - Kohl's

ML DREAM

What product you would like...

The image shows a Google search interface for the query "shoes". The search bar at the top contains the word "shoes" and shows a search count of "About 1,810,000,000 results (1.10 seconds)". Below the search bar are navigation tabs for "All", "Shopping", "Maps", "Images", "News", "More", "Settings", and "Tools".

The search results are divided into two main sections. The left section, highlighted with a red border, contains four organic search results:

- Shoes at Zappos.com**: An advertisement for Zappos.com with a 4.7-star rating. It highlights "Fast, Free Shipping & Free 365 Day Returns on Huge Selection of Shoes!" and lists brands like Birkenstock, Nike, Converse, New Balance, Born, and Frye. Types include Sneakers, Slippers, Heels, Boots, Flats, and Running Shoes.
- Womens Shoes at Macy's - Save 40-60% on Cyber Monday - macys.com**: An advertisement for Macy's with a 4.2-star rating. It promotes "Cyber Monday Ends 11/30, Deals Going Fast + Extra 20% Off!" and lists styles like Boots, Wedges, Loafers, Flats, Slippers, Sneakers, Pumps, Espadrilles, and Booties. Ratings for Returns, Product quality, Shipping, Service, and Selection are all 8.5/10.
- Converse® Official Site - Converse.com**: An advertisement for Converse.com with a 4.2-star rating. It promotes "Full Converse Collection. Shop Our New Designs for Men, Women & Kids." and lists types like Men's Sneakers, Women's Sneakers, Kid's Sneakers, and Custom Sneakers.
- Brand Name Shoes for Less - Up to 70% Off - zulily.com**: An advertisement for Zulily.com with a 4.2-star rating. It promotes "Exclusive Deals On All Shoes, Socks, & More at Zulily. Shop & Save Today!"

The right section, also highlighted with a red border, features a sponsored product listing:

- Shop on Google**: A sponsored listing for Adidas NEO Baseline Women's Dicast-Leather Sneakers, Size: 8, White, priced at \$34.99 from Kohl's. The listing includes a small image of the shoe and a "Sponsored" label.

At the bottom of the page, there is a link for "Men and Womens shoes, Shipped Free | Zappos.com" with a URL to www.zappos.com/shoes. Below this link are several category links: "Boots Sneakers & Athletic Heels Flats Sandals View All...", "Sneakers & Athletic Boots Oxford's Loafers Sandals View All...", "Sneakers and Athletic Boots Slippers Flats Sandals view all...", "Sandals", "Womens Shoes", "Sneakers & Athletic Shoes", and "Popular Mens Shoe Styles".

ML DREAM

What news would you prefer to read...



ML DREAM

Find the best job for you...

The screenshot shows the LinkedIn job search interface. At the top, the LinkedIn logo is on the left, and navigation links for 'What is LinkedIn?', 'Join Today', and 'Sign In' are on the right. Below the navigation bar is a search bar containing 'Data Scientist' and 'United States', with a 'Find jobs' button to the right. On the left side, there are three filter panels: 'Get alerts for this search' with an email input field and a 'Create job alert' button; 'Location' with a list of cities and their job counts; and 'Company' with a list of companies and their job counts. The main content area displays a list of job results, each with a company logo, job title, location, description, and an 'Apply with Profile' button. The results are sorted by 'Relevance'.

LinkedIn

What is LinkedIn? Join Today Sign In

Data Scientist United States Find jobs

Get alerts for this search
We'll email you new jobs as they become available

Email address

Create job alert

Location

- New York, New York (621)
- San Francisco, California (575)
- Seattle, Washington (318)
- Chicago, Illinois (291)
- Atlanta, Georgia (181)

Company

- Maverick Trading (339)
- Amazon (265)
- Deloitte (235)
- CyberCoders (165)
- Jobspring Partners (145)

+ Add

Date Posted

8,715 Data Scientist jobs in United States sort by: Relevance

Data Scientist 13d
LeadGenius
San Francisco Bay Area
We are looking for a seasoned Data Scientist/Machine Learning engineer to build the next generation mission critical data platform. Solid engineering and coding skills.
Apply with Profile

Data Scientist 8d
feedzai
Atlanta, Georgia
...client's data feeds Work with the the client to explore their data and better understand it Work...

Data Scientist 17d
Jetlore
Sunnyvale, California
We are looking for an exceptional data scientist who is excited to work on challenging problems involving massive amount of data. Ping-pong skills is a plus!

Data Scientist 5d
Covestro
Greater Pittsburgh Area
Covestro is in search of a Data Scientist... and data analysis to help influence changes...

ML DREAM

- For every user predict: Ads, products, news, ...
- Have tons of data to learn this task well
- Have right models that can learn from all this data

ML DREAM

- For every user predict: Ads, products, news, ...
- Have tons of data to learn this task well
- Have right models that can learn from all this data

With Big Data comes Bigger Responsibilities ...

IS ML FAIR, IMPARTIAL?

Google

Women less likely to be shown ads for high-paid jobs on Google, study shows

Automated testing and analysis of company's advertising system reveals male job seekers are shown far more adverts for high-paying executive jobs

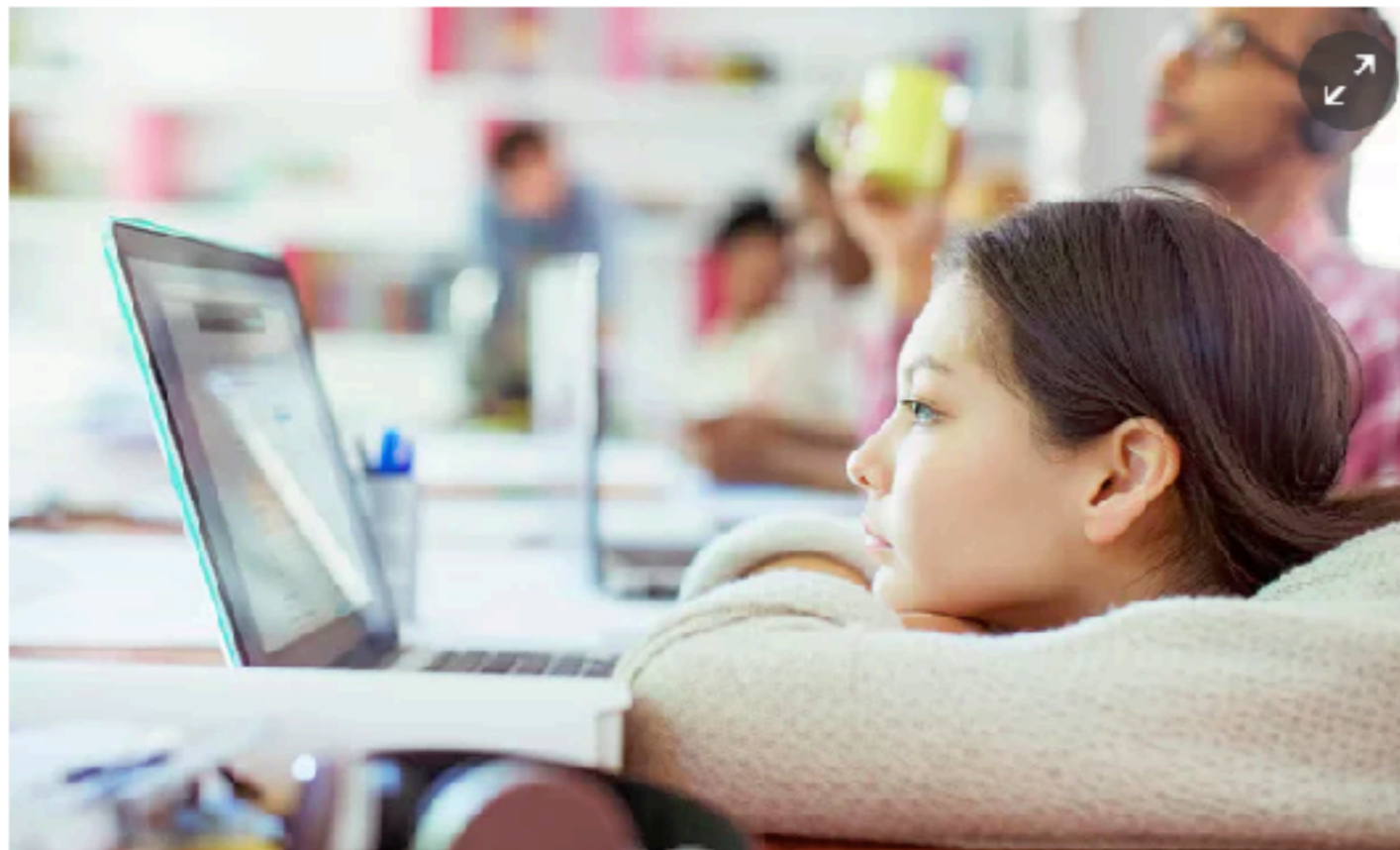
Samuel Gibbs

Wednesday 8 July 2015
06:29 EDT



This article is 1 year old

1120 140



One experiment showed that Google displayed adverts for a career coaching service for executive jobs 1,857 times to the male group and only 318 times to the female group. Photograph: Alamy

Female job seekers are much less likely to be shown adverts on [Google](#) for highly paid jobs than men, researchers have found.

Advertisement



In the spirit of Giving Tuesday, we're donating our ads to charity today

IS ML FAIR, IMPARTIAL?

Prediction Fails Differently for Black Defendants

	WHITE	AFRICAN AMERICAN
Labeled Higher Risk, But Didn't Re-Offend	23.5%	44.9%
Labeled Lower Risk, Yet Did Re-Offend	47.7%	28.0%

Overall, Northpointe's assessment tool correctly predicts recidivism 61 percent of the time. But blacks are almost twice as likely as whites to be labeled a higher risk but not actually re-offend. It makes the opposite mistake among whites: They are much more likely than blacks to be labeled lower risk but go on to commit other crimes. (Source: ProPublica analysis of data from Broward County, Fla.)

Can we make ML Fair?

- These are machine learning algorithms that learn to predict automatically
- They are not designed to be unfair
- Why is this happening?
- How do we fix them?

WHY NOW?

WHY NOW?



Loads of data collected everywhere!

WHY NOW?



Machine Learning

galvanize

 Startup.ML

coursera



datascience@berkeley

 The Data Incubator

METIS

WHY NOW?



Machine Learning

galvanize

 Startup.ML

courseera



datascience@berkeley

 The Data Incubator

METIS

Raise in number of Data Scientists!

WHY IS ML UNFAIR?

the algorithms in themselves are neutral. “This program had absolutely nothing to do with race... but multi-variable equations,”

WHY IS ML UNFAIR?

- Data collection, labeling etc. can have unintentional biases
 - We learn from past data, historic biases
- Data in itself nor algorithms explicitly know of social inequities

FAIRNESS THROUGH BLINDNESS?

- Ignore all protected attributes.
Eg. Don't look at race, gender etc.

FAIRNESS THROUGH BLINDNESS?

- Ignore all protected attributes.
Eg. Don't look at race, gender etc.
- **Problem: You don't need to look to be able to predict**

FAIRNESS THROUGH BLINDNESS?

- Ignore all protected attributes.
Eg. Don't look at race, gender etc.
- **Problem: You don't need to look to be able to predict**

Eg. User visits "www.artofmanliness.com"
...highly likely to be male

EG. REAL VS FAKE NAMES

- Biases are often not intentional ...

EG. REAL VS FAKE NAMES



- Biases are often not intentional ...
- Most training examples standard white American names: James, John, Robert, Jennifer, Michael, ...
- Ethnic names often unique, much fewer training examples

EG. REAL VS FAKE NAMES

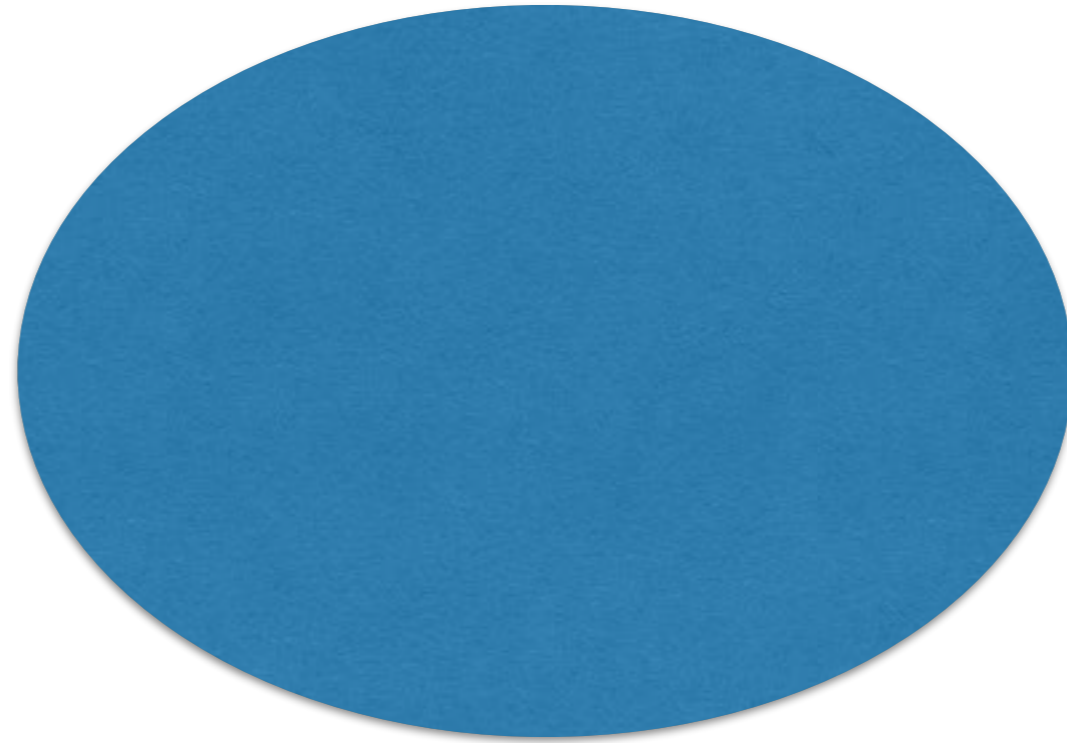


- Biases are often not intentional ...
- Most training examples standard white American names: James, John, Robert, Jennifer, Michael, ...
- Ethnic names often unique, much fewer training examples

Most ML models aim for accuracy for the majority at the expense of mistakes on the smaller protected class

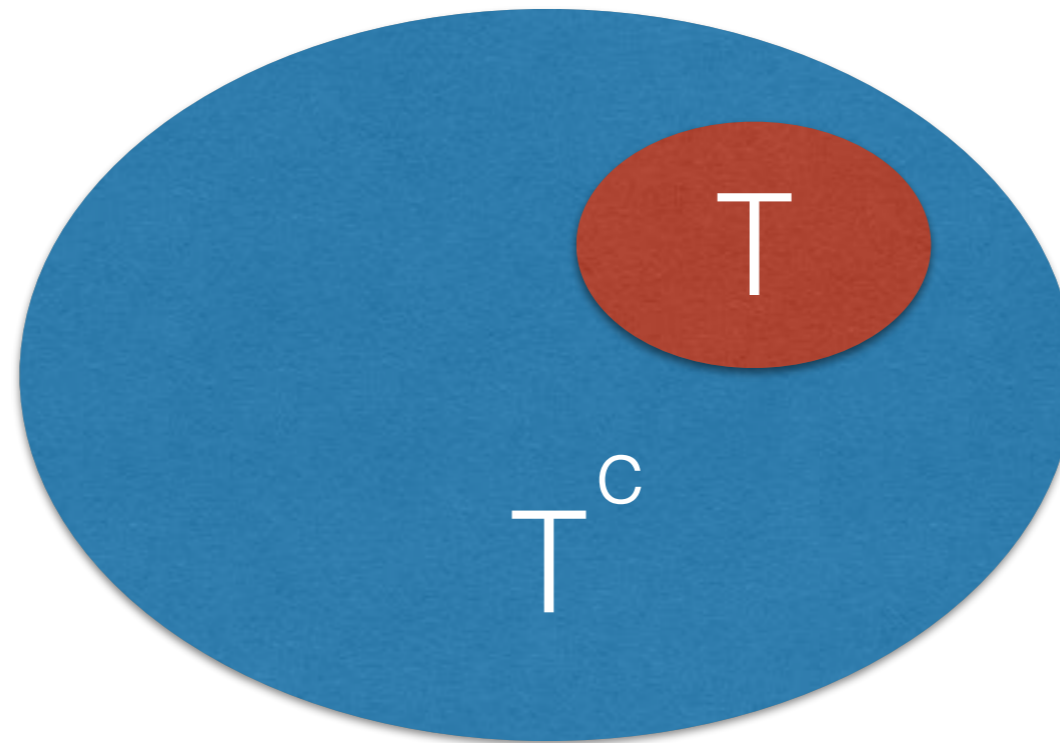
FAIRNESS THROUGH AWARENESS

FAIRNESS THROUGH AWARENESS



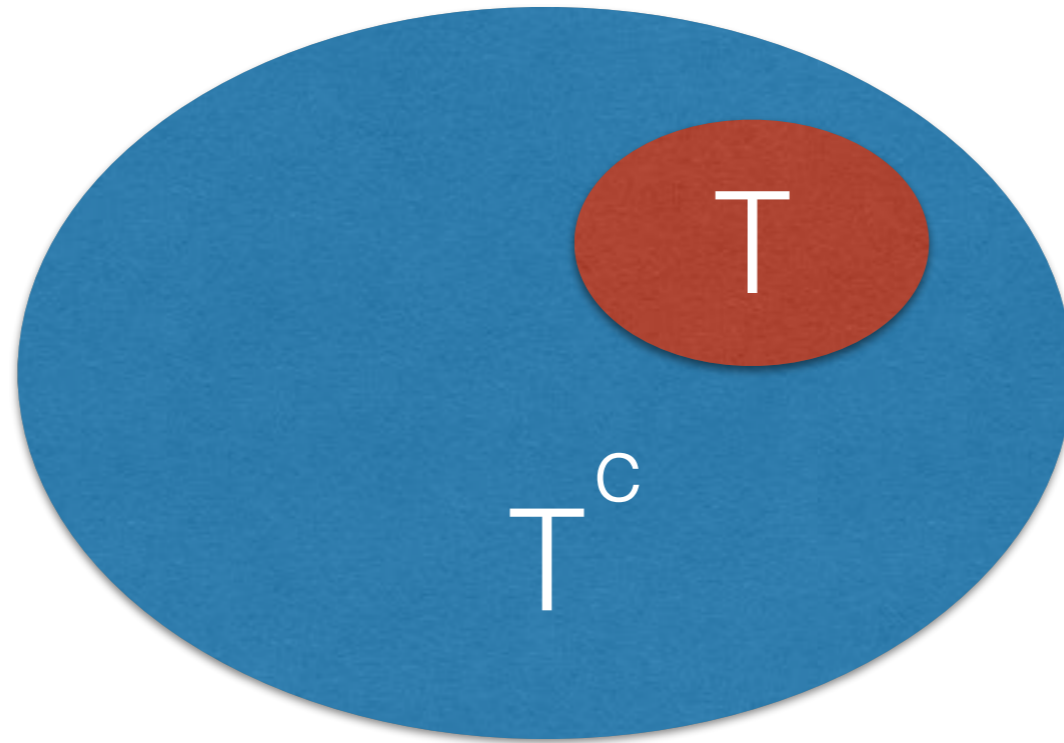
Population

FAIRNESS THROUGH AWARENESS



Population

FAIRNESS THROUGH AWARENESS

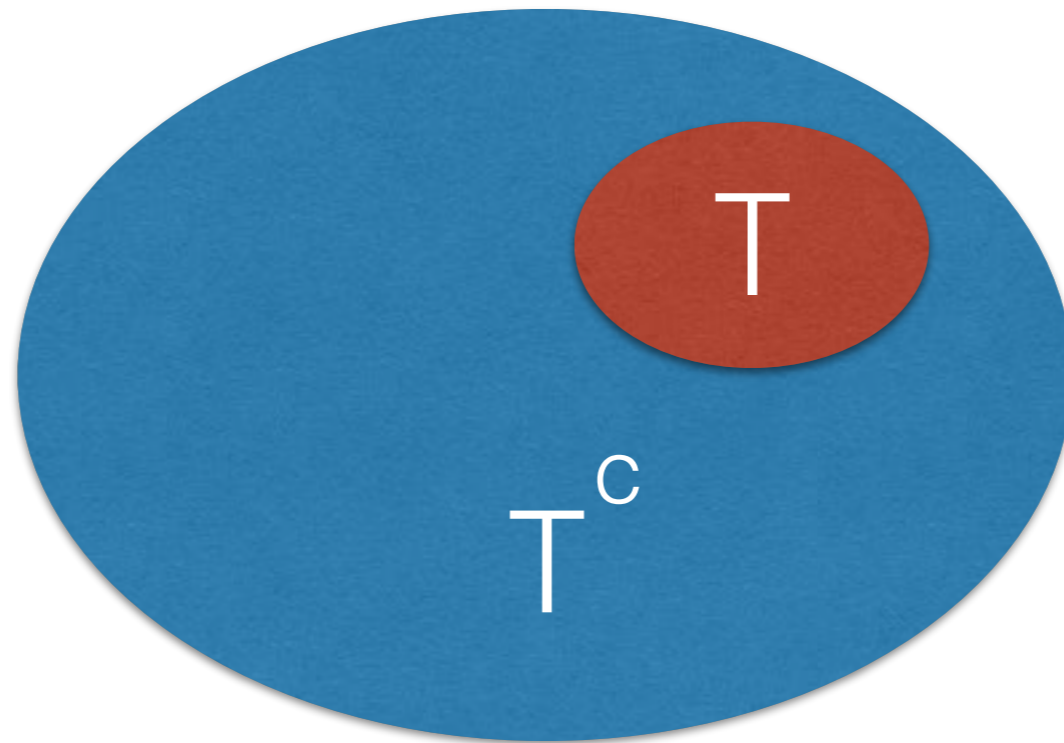


Population

T : Protected subset

T^c : Rest of the population

FAIRNESS THROUGH AWARENESS



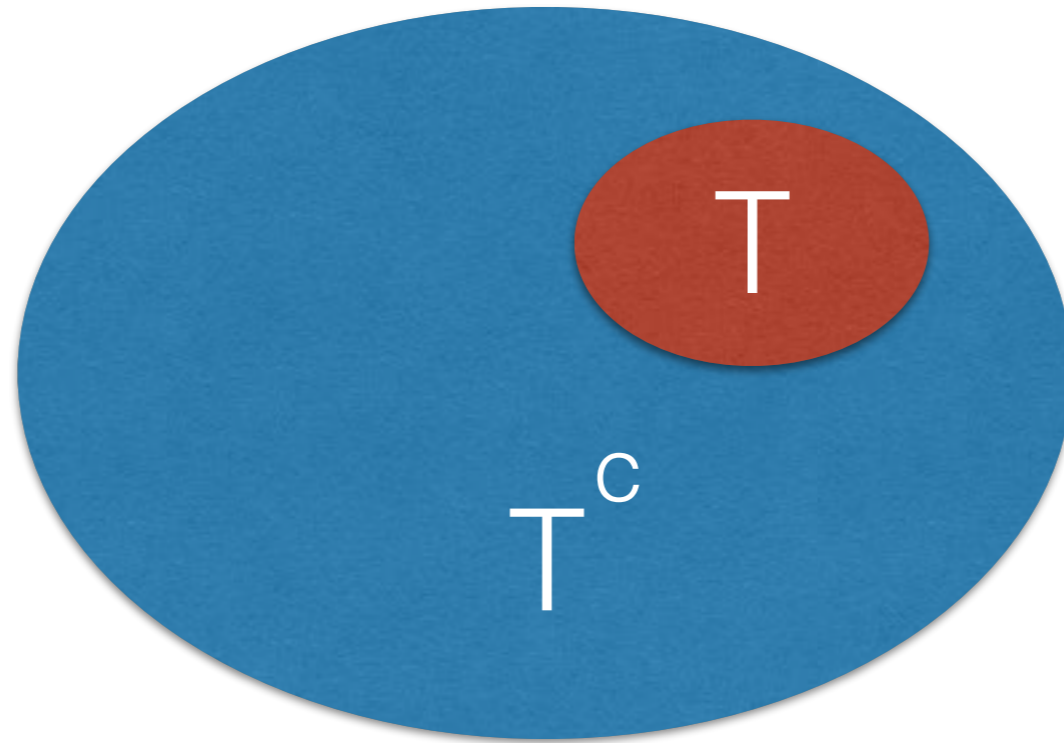
Population

T : Protected subset

T^c : Rest of the population

$$P(\text{Outcome}|T) \approx P(\text{Outcome}|T^c)$$

FAIRNESS THROUGH AWARENESS



Population

Eg. Fraction of people shown high paying jobs in T and in T^c is equal

EG. FAIR K-MEANS CLUSTERING (very naive)

EG. FAIR K-MEANS CLUSTERING (very naive)

$$\text{Objective} = \sum_{j=1}^K \sum_{t \in C_j} \|\mathbf{x}_t - \mathbf{r}_j\|_2^2$$

$$\text{where } \mathbf{r}_j = \frac{1}{|C_j|} \sum_{\mathbf{x}_t \in C_j} \mathbf{x}_t$$

EG. FAIR K-MEANS CLUSTERING (very naive)

$$\text{Objective} = \sum_{j=1}^K \sum_{t \in C_j} \|\mathbf{x}_t - \mathbf{r}_j\|_2^2$$

$$\text{where } \mathbf{r}_j = \frac{1}{|C_j|} \sum_{\mathbf{x}_t \in C_j} \mathbf{x}_t$$

$$\text{Fairness constraints: } \forall j \in [K], \quad \sum_{t: c_t=j} \mathbf{1}_{x_t \in T} = \sum_{t: c_t=j} \mathbf{1}_{x_t \notin T}$$

EG. FAIR K-MEANS CLUSTERING (very naive)

$$\text{Objective} = \sum_{j=1}^K \sum_{t \in C_j} \|\mathbf{x}_t - \mathbf{r}_j\|_2^2$$

$$\text{where } \mathbf{r}_j = \frac{1}{|C_j|} \sum_{\mathbf{x}_t \in C_j} \mathbf{x}_t$$

$$\text{Fairness constraints: } \forall j \in [K], \quad \sum_{t: c_t=j} \mathbf{1}_{x_t \in T} = \sum_{t: c_t=j} \mathbf{1}_{x_t \notin T}$$

Number of protected in cluster j = Number of unprotected in cluster j

FAIR CLASSIFICATION

A view from a mile above:

FAIR CLASSIFICATION

A view from a mile above:

Minimize Classification objective
(or whatever other surrogate loss you use usually)

FAIR CLASSIFICATION

A view from a mile above:

Minimize Classification objective
(or whatever other surrogate loss you use usually)

Added Constraint: subject to proportion of labels in each class being same for protected and unprotected population

FAIR CLASSIFICATION

FAIR CLASSIFICATION

- Is this good enough?

FAIR CLASSIFICATION

- Is this good enough?
- Say there is this algorithm to select people to invite to apply for this **exclusive, credit card with high annual fee**

FAIR CLASSIFICATION

- Is this good enough?
- Say there is this algorithm to select people to invite to apply for this **exclusive, credit card with high annual fee**
- One simple way to satisfy the fairness constraint:

FAIR CLASSIFICATION

- Is this good enough?
- Say there is this algorithm to select people to invite to apply for this **exclusive, credit card with high annual fee**
- One simple way to satisfy the fairness constraint:
 - Make offer to higher income people in the unprotected class

FAIR CLASSIFICATION

- Is this good enough?
- Say there is this algorithm to select people to invite to apply for this **exclusive, credit card with high annual fee**
- One simple way to satisfy the fairness constraint:
 - Make offer to higher income people in the unprotected class
 - Make offer to lower income people in protected class (in same proportion)

FAIR CLASSIFICATION

- Is this good enough?
- Say there is this algorithm to select people to invite to apply for this **exclusive, credit card with high annual fee**
- One simple way to satisfy the fairness constraint:
 - Make offer to higher income people in the unprotected class
 - Make offer to lower income people in protected class (in same proportion)

NOT REALLY FAIR!

FAIR CLASSIFICATION

A view from a mile above:

Minimize Classification objective
(or whatever other surrogate loss you use usually)

FAIR CLASSIFICATION

A view from a mile above:

Minimize Classification objective
(or whatever other surrogate loss you use usually)

Added Constraint: subject to proportion of labels for
similar instances
in each class being same for protected
and unprotected population

On to the next social issue...

ML DREAM

What news would user prefer to read...



ML DREAM

ML DREAM

- Just because we can predict, should we?

ML DREAM

- Just because we can predict, should we?
- Say we have a fair, unbiased algorithm for prediction

ML DREAM

- Just because we can predict, should we?
- Say we have a fair, unbiased algorithm for prediction
- Can there be other issues?

EXTREMIZING EFFECT OF ML

EXTREMIZING EFFECT OF ML



EXTREMIZING EFFECT OF ML



User 1

User 2

EXTREMIZING EFFECT OF ML



User 1

Apples are extremely rich in important antioxidants, flavanoids, and dietary fiber. The phytonutrients and antioxidants in **apples** may help reduce the risk of developing cancer, hypertension, diabetes, and heart



User 2

For fewer calories per fruit, **oranges** have higher levels of Vitamin C, folate, potassium, and protein.

EXTREMIZING EFFECT OF ML



User 1

Apples are extremely rich in important antioxidants, flavanoids, and dietary fiber. The phytonutrients and antioxidants in **apples** may help reduce the risk of developing cancer, hypertension, diabetes, and heart

TOP 10 Health Benefits of Apples

www.herbs-info.com

1. Cancer Prevention
2. Antioxidant Activity
3. Antihyperglycemic
4. Anti-diabetes
5. Cardiovascular Protection
6. Cholesterol Reduction
7. Anti-asthma
8. Weight Reduction

Apples



User 2

For fewer calories per fruit, **oranges** have higher levels of Vitamin C, folate, potassium, and protein.

The Health Benefits of Oranges

- Packed with fiber to promote healthy digestion
- Full of folate to help the body form red blood cells
- A good source of immune-boosting vitamin C
- Contains potassium to ensure a healthy heart
- Keeps vision clear and eyes healthy with its content of vitamin A
- A great source of calcium for healthy and strong bones



EXTREMIZING EFFECT OF ML



User 1

Apples are extremely rich in important antioxidants, flavanoids, and dietary fiber. The phytonutrients and antioxidants in **apples** may help reduce the risk of developing cancer, hypertension, diabetes, and heart

TOP 10 Health Benefits of Apples

- 1. Cancer Prevention
- 2. Antioxidant Activity
- 3. Antihyperglycemic
- 4. Anti-diabetes
- 5. Cardiovascular Protection
- 6. Cholesterol Reduction
- 7. Anti-asthma
- 8. Weight Reduction



User 2

For fewer calories per fruit, **oranges** have higher levels of Vitamin C, folate, potassium, and protein.

The Health Benefits of Oranges

- Packed with fiber to promote healthy digestion
- Full of folate to help the body form red blood cells
- A good source of immune-boosting vitamin C
- Contains potassium to ensure a healthy heart
- Keeps vision clear and eyes healthy with its content of vitamin A
- A great source of calcium for healthy and strong bones



EXTREMIZING EFFECT OF ML



Vs



User 1

Apples are extremely rich in important antioxidants, flavanoids, and dietary fiber. The phytonutrients and antioxidants in **apples** may help reduce the risk of developing cancer, hypertension, diabetes, and heart



User 2

For fewer calories per fruit, **oranges** have higher levels of Vitamin C, folate, potassium, and protein.

The Health Benefits of Oranges

- Packed with fiber to promote healthy digestion
- Full of folate to help the body form red blood cells
- A good source of immune-boosting vitamin C
- Contains potassium to ensure a healthy heart
- Keeps vision clear and eyes healthy with its content of vitamin A
- A great source of calcium for healthy and strong bones



EXTREMIZING EFFECT OF ML

EXTREMIZING EFFECT OF ML

- Possible fixes:

EXTREMIZING EFFECT OF ML

- Possible fixes:
 - Mix up articles: provide both sides of a story

EXTREMIZING EFFECT OF ML

- Possible fixes:
 - Mix up articles: provide both sides of a story
 - what if there are multiple sides and worse yet you can't identify the sides clearly

EXTREMIZING EFFECT OF ML

- Possible fixes:
 - Mix up articles: provide both sides of a story
 - what if there are multiple sides and worse yet you can't identify the sides clearly
 - Mix up user profiles from time to time

EXTREMIZING EFFECT OF ML

- Possible fixes:
 - Mix up articles: provide both sides of a story
 - what if there are multiple sides and worse yet you can't identify the sides clearly
 - Mix up user profiles from time to time
 - Learn a canonical recommendation model for an “average” user

EXTREMIZING EFFECT OF ML

- Possible fixes:
 - Mix up articles: provide both sides of a story
 - what if there are multiple sides and worse yet you can't identify the sides clearly
 - Mix up user profiles from time to time
 - Learn a canonical recommendation model for an “average” user
 - every user is mix of the individual and the canonical user

EXTREMIZING EFFECT OF ML

- Issue: assimilation bias
 - Users tend to view opposing arguments more skeptically
 - If we show a biased user one For and one Against article of same strength, they tend to get more biased towards their (“For”) view

EXTREMIZING EFFECT OF ML

EXTREMIZING EFFECT OF ML

- Possible approach:

EXTREMIZING EFFECT OF ML

- Possible approach:
 - Use external or other information like the social networks graph. Eg. Typically ideologically similar users form tight knit groups on social networks

EXTREMIZING EFFECT OF ML

- Possible approach:
 - Use external or other information like the social networks graph. Eg. Typically ideologically similar users form tight knit groups on social networks
 - Use social network to de-bias articles provided

EXTREMIZING EFFECT OF ML

- Possible approach:
 - Use external or other information like the social networks graph. Eg. Typically ideologically similar users form tight knit groups on social networks
 - Use social network to de-bias articles provided
 - Eg. recommend articles on which your more moderate friends disagree with you on

EXTREMIZING EFFECT OF ML

- Possible approach:
 - Use external or other information like the social networks graph. Eg. Typically ideologically similar users form tight knit groups on social networks
 - Use social network to de-bias articles provided
 - Eg. recommend articles on which your more moderate friends disagree with you on
 - This is a completely open topic ...

EXTREMIZING EFFECT OF ML

- Possible approach:
 - Use external or other information like the social networks graph. Eg. Typically ideologically similar users form tight knit groups on social networks
 - Use social network to de-bias articles provided
 - Eg. recommend articles on which your more moderate friends disagree with you on
 - This is a completely open topic ...
 - But a very important one ...

TRANSPARENCY IN ML

- Another issue: ML methods are complex and we don't understand semantic meaning

TRANSPARENCY IN ML

- Another issue: ML methods are complex and we don't understand semantic meaning
- We need transparency of method for accountability

TRANSPARENCY IN ML

- Another issue: ML methods are complex and we don't understand semantic meaning
- We need transparency of method for accountability
- Transparency via interpretability.
 - Provide explanation for each decision
 - What makes an instance a negative instance according to the algorithm

Fairness, Accountability, and Transparency in Machine Learning

DATA
TRANSPARENCY
LAB

DTL Conferences ▾

Become a sponsor

Grants Program ▾

Technical Program ▾

About ▾

Up to Date

Workshop ▾

ES

Contact

Newsletter



DTL CONFERENCES

DTL 2017

Barcelona **December 11-13th**



Time for General Questions