

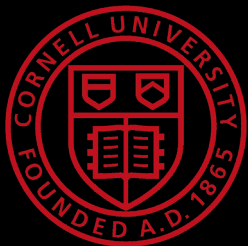
# Bias and Fairness in Learning Systems

CS 7792 - Spring 2020

Thorsten Joachims

Department of Computer Science & Department of Information Science

Cornell University



# Outline of Today

- Introduction
  - Thorsten Joachims
- Motivation of Class Topics
- Administrivia
  - Goals for the Class
  - Pre-Requisites
  - Credit Options and Format
  - Course Material
  - Contact Info

# Bias and Fairness in ML

Technology **BBC**

## Google hit with record EU fine over Shopping service

By Leo Kelion  
Technology desk editor

## Wenn Maschinen kalt entscheiden

Algorithmen urteilen über Arbeitslose, Straftäter und Job-Bewerber. Eine Kommission der Bundesregierung will ihre Macht nun bändigen. Kann das gelingen?

Eine Analyse von **Ann-Kathrin Nezik** ZEITUNG ONLINE

23. Oktober 2019, 16:46 Uhr / Editiert am 24. Oktober 2019, 15:40 Uhr / DIE ZEIT Nr. 44/2019, 24. Oktober 2019 / [9 Kommentare](#)



## Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

by *Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica*

May 23, 2016

BUSINESS NEWS OCTOBER 9, 2018 / 11:12 PM / A YEAR AGO



## Amazon scraps secret AI recruiting tool that showed bias against women

Jeffrey Dastin

8 MIN READ



# Two Motivating Examples

→ Example 1:

Rankings in Recommender Systems

Example 2:

College Admission

# Ranking in Online Systems

Ranking function  $\pi$  that ranks items for context  $x$ .  
→ Learning-to-Rank

svm at DuckDuckGo

duckduckgo.com/?q=svm&ts=s&ia=about

svm

Web Images Videos News Maps Meanings Stock Settings

All Regions Safe Search Moderate Any Time

### Support-vector machine

In machine learning, support-vector machines are supervised learning models, with associated learning algorithms that analyze data used for classification and regression analysis. Wikipedia

Support-vector machine - Wikipedia

https://en.wikipedia.org/wiki/Support\_vector\_machine

In machine learning, support-vector machines (SVMs), also support-vector networks) are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis.

SVM - Gas Gift Card & Promotion Management

https://www.svmcards.com

SVM is a leading provider of gift card marketing, management & customization services. Find over 250 brands promotional gift card solutions for gas, retail & services

Silvercorp Metals Inc. (SVM) - Yahoo Finance

https://finance.yahoo.com/quote/SVM/

Find the latest Silvercorp Metals Inc. (SVM) stock quote, history, news and other vital information to help you with your stock trading and investing.

Introduction to Support Vector Machine(SVM) Dimensionless ...

https://dimensionless.in/introduction-to-svm/

A Support Vector Machine(SVM) is yet another supervised machine learning algorithm. It

metal table legs | Etsy

etsy.com/search?q=metal%20table%20legs

## Etsy

metal table legs

metal dining table legs metal coffee table legs metal hairpin table legs table legs

All categories Home & Living Craft Supplies & Tools Art & Collectibles Paper & Party Supplies

Special offers

FREE shipping On sale

Ready to ship in 1 business day 1-3 business days

Price (\$)

Any price Under \$100 \$100 to \$500 \$500 to \$1,000 Over \$1,000 Custom

Low to High

Style

All categories > "metal table legs" (5,758 Results)

Ad

Heavy Duty Metal Table Legs... ModernUrbanMetals

4.9 (4,165)

\$15.74 \$22.48 (30% off)

FREE shipping

Ad

2 Pack U Shape Legs (1.5" W)... CiomkaCustomRustics

4.9 (1,960)

\$16.20 \$48.00 (70% off)

FREE shipping

Ad

Heavy Duty Metal Table Legs... ModernUrbanMetals

4.9 (4,165)

\$15.74 \$22.48 (30% off)

FREE shipping

Ad

2 Pack - (2" Wide - 1/4" Thick)... CiomkaCustomRustics

4.9 (1,960)

\$32.40 \$96.00 (70% off)

FREE shipping

Netflix

netflix.com/browse

## NETFLIX

Durchsuchen

TOAST LONDON

SPECIAL COLLECTIONS

HAPPYISH

Detectiv-Dramserien

CRIMINAL MINDS

hinterland

REPUBLIC DOYLE

NETFLIX ORIGINALS

THE GREAT BRITISH BAKING SHOW

THE LAST KINGDOM

The Knives in Method

HuffPost - Breaking News

huffpost.com/?guccounter=1&guce\_referrer=aHR0LmML...

## HUFFPOST

Log In Join HuffPost Plus

TOP STORIES

Pelosi Announces House Will Vote On Impeachment Process After Weeks of GOP Protests

Katie Hill Speaks Out On Resignation: 'I'm Hurt, I'm Angry'

Court Strikes Down North Carolina Congressional Map

Democratic Former Sen. Kay Hagan Dies At 66

Trump Publicly Revels In Death And Spectacle: 'Boom Boom Boom!'

Palbear Who Snubbed Mitch McConnell At Elijah Cummings' Memorial Breaks Silence

U.S. LANDS CHIEF WROTE ANTI-ENVIRONMENTAL SCREEDS

Trump's Public Lands Chief Wrote For A Cult Extremist's Magazine

By Chris D'Angelo and Alexander C. Kaufman

UKRAINE CALL WITNESS NO-SHOWS IMPEACHMENT HEARING

Former White House Aide Won't Show For Scheduled Impeachment

Writing for info.olegromed.com

# Two-Sided Market

## Online Retail

- Utility to Users:  
Customers find products they want
- Utility to Items:  
Sellers get revenue

The screenshot shows a web browser displaying the Etsy search results for 'metal table legs'. The browser's address bar shows the URL 'etsy.com/search?q=metal%20table%20legs'. The Etsy logo is visible in the top left, and a 'Sign in' button is in the top right. Below the search bar, there are filter tabs for 'metal dining table legs', 'metal coffee table legs', 'metal hairpin table legs', and 'table legs'. The search results are displayed in a grid format, showing four items. Each item includes a product image, a title, the seller's name, a star rating with the number of reviews, the current price, the original price, and the discount percentage. The first two items are 'Heavy Duty Metal Table Legs...' and '2 Pack U Shape Legs (1.5" Wi...' by ModernUrbanMetals and CsonkasCustomRustics respectively. The last two items are 'Heavy Duty Metal Table Legs...' and '2 Pack - (2" Wide - 1/4" Thick ...' by ModernUrbanMetals and CsonkasCustomRustics respectively. The page also features a left sidebar with filters for 'All categories', 'Special offers', 'Ready to ship in', 'Price (\$)', and 'Style'.

metal table legs | Etsy

etsy.com/search?q=metal%20table%20legs

Apps FamilyCalendar Thermostat Ooma COECIS-IT CMS UGrad Other bookmarks

Etsy Sign in

metal table legs

metal dining table legs metal coffee table legs metal hairpin table legs table legs

All categories

Home & Living  
Craft Supplies & Tools  
Art & Collectibles  
Paper & Party Supplies  
+ Show more

Special offers

FREE shipping  
 On sale

Ready to ship in

1 business day  
 1 - 3 business days

Price (\$)

Any price  
 Under \$100  
 \$100 to \$500  
 \$500 to \$1,000  
 Over \$1,000  
 Custom

Low to High >

Style

Ad

Heavy Duty Metal Table Legs...  
ModernUrbanMetals  
★★★★★ (4,165)  
\$15.74 \$22.48 (30% off)  
FREE shipping

Ad

2 Pack U Shape Legs (1.5" Wi...  
CsonkasCustomRustics  
★★★★★ (1,960)  
\$16.20 \$18.00 (10% off)  
FREE shipping

Ad

Heavy Duty Metal Table Legs...  
ModernUrbanMetals  
★★★★★ (4,165)  
\$15.74 \$22.48 (30% off)  
FREE shipping

Ad

2 Pack - (2" Wide - 1/4" Thick ...  
CsonkasCustomRustics  
★★★★★ (1,960)  
\$32.40 \$36.00 (10% off)  
FREE shipping

# Two-Sided Market

## Music Streaming

- Utility to Users:  
Customers find music they enjoy
- Utility to Items:  
Artists get streaming revenue

The screenshot displays the Spotify web interface. At the top, the browser address bar shows 'open.spotify.com/browse/featured'. The main content area is divided into several sections:

- Recently played:** A row of three album covers: 'Kid's Car Mix', 'Marmor, Stein & Eisen Bricht' by Drafi Deutscher, and 'KIDZ BOP 39 (Deluxe Edition)' by Kidz Bop Kids.
- Your heavy rotation:** A section titled 'The music you've had on repeat this month.' featuring three items: 'Slaapliedjes', 'The Greatest Showman (Original...)' by Various artists, and 'Annie 2014'.

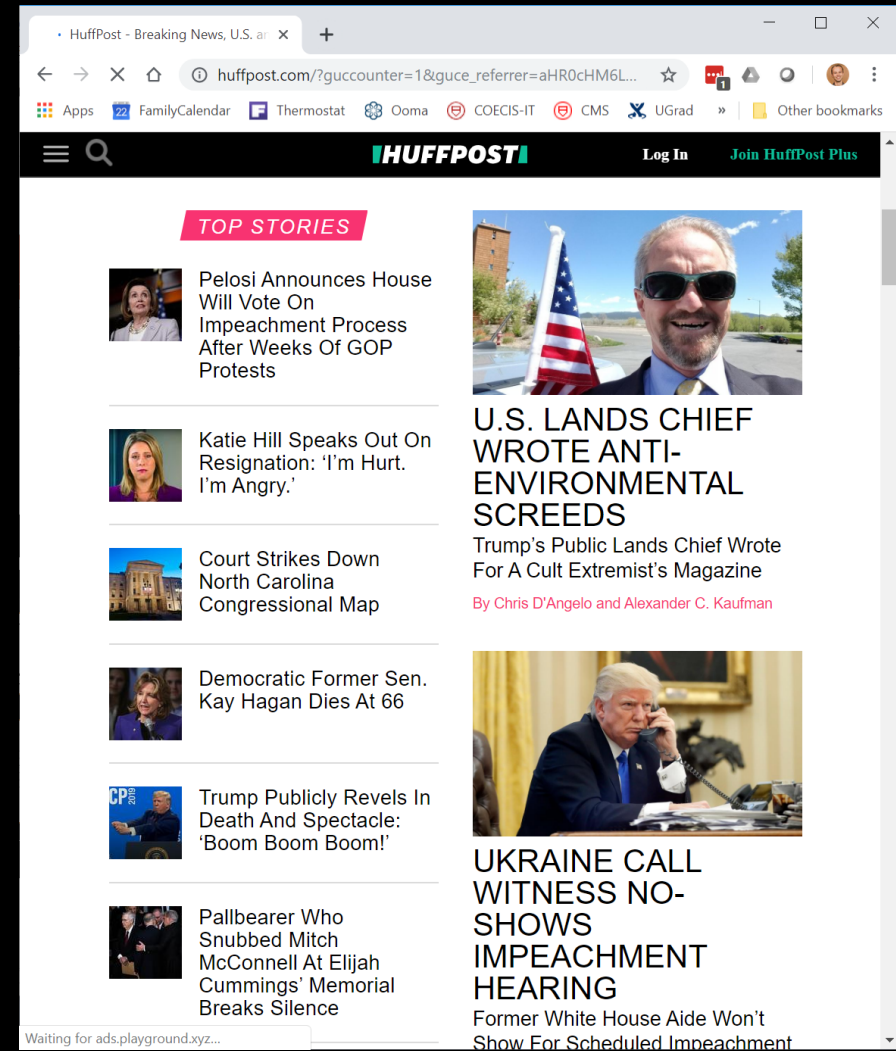
The left sidebar contains navigation options: Home, Search, Your Library, and a 'PLAYLISTS' section with a 'Create Playlist' button. Below this, a list of playlists is visible, including 'Localify - Ithaca', 'Nena - Nena 40 - Das...', 'Victor Davich - 8 Minu...', 'Hold Hope', 'Annie 2014', 'Ukulele beats', 'ukulele', 'Calming Christian', 'Spread the Gospel', and 'Rhythm & Praise'. At the bottom of the sidebar, there is an 'Install App' button and a profile icon for 'Thorsten Joachims'.

The bottom of the image shows a music player interface with a play button, a progress bar at 0:00, and a volume control icon.

# Two-Sided Market

## News

- Utility to Users:  
Readers find relevant articles
- Utility to Items:  
Writers get their voice out (and ad revenue)



The screenshot shows the HuffPost website interface. At the top, there's a navigation bar with the HuffPost logo, 'Log In', and 'Join HuffPost Plus' links. Below this is a 'TOP STORIES' section with a pink header. The stories are listed in two columns. The left column contains five articles, each with a small thumbnail image and a headline. The right column features a larger article with a prominent image of a man in sunglasses and a headline about the U.S. Lands Chief. At the bottom of the page, there's a small text box that says 'Waiting for ads.playground.xyz...'.


HuffPost - Breaking News, U.S. a | x +


huffpost.com/?guccounter=1&guce\_referrer=aHR0cHM6L... ☆


Apps FamilyCalendar Thermostat Ooma COECIS-IT CMS UGrad Other bookmarks


HUFFPOST Log In Join HuffPost Plus


**TOP STORIES**


 Pelosi Announces House Will Vote On Impeachment Process After Weeks Of GOP Protests


 Katie Hill Speaks Out On Resignation: 'I'm Hurt. I'm Angry.'


 Court Strikes Down North Carolina Congressional Map

 Democratic Former Sen. Kay Hagan Dies At 66

 Trump Publicly Reveals In Death And Spectacle: 'Boom Boom Boom!'

 Pallbearer Who Snubbed Mitch McConnell At Elijah Cummings' Memorial Breaks Silence

 **U.S. LANDS CHIEF WROTE ANTI-ENVIRONMENTAL SCREEDS**  
Trump's Public Lands Chief Wrote For A Cult Extremist's Magazine  
By Chris D'Angelo and Alexander C. Kaufman

 **UKRAINE CALL WITNESS NO-SHOWS IMPEACHMENT HEARING**  
Former White House Aide Won't Show For Scheduled Impeachment

Waiting for ads.playground.xyz...



# Are Conventional Methods Fair?

Probability Ranking Principle:

- Rank documents by probability of relevance  $\rightarrow y^*$  [Robertson, 1977]
- For virtually any measure  $U$  of ranking quality

$$y^* := \operatorname{argmax}_y [U(y|x)]$$

- Are rankings fair/desirable?

Top News Stories		
Rank	Item	P(read)
1	Times 1	50.99
2	Times 2	50.98
3	Times 3	50.97
...	...	...
100	Post 1	49.99
101	Post 2	49.98
102	Post 3	49.97
...	...	...

# Fairness in Ranking

## Endogenous Factors

How to allocate exposure based on merit in order to

- Satisfy legal requirements
- Shape marketplace dynamics (e.g. Spotify, superstar economics)
- Spam, Polarization

## Exogenous Factors

How to estimate merit without biases like

- Position bias
- Trust bias
- Uncertainty bias
- Historical actions
- Stereotypes

# Two Motivating Examples

Example 1:

Rankings in Recommender Systems

→ Example 2:

College Admission

# College Admission

## Problem

- 45,000 applications → 6,000 admissions
- Manual reading and evaluation by humans
- Noise, variability by reviewer, cost

How can this be (partially) automated to

- Make better decisions
- Reduce cost

# Some Strawmen Proposals

Why not do the following?

- Sort students by high-school GPA.
- Sort students by SAT score.
- Sort students by SAT score for each legally protected class (race, color, religion or creed, national origin or ancestry, sex, age, disability, veteran status, genetic information, citizenship).

# Broader Picture

Balance cost/benefit to

- Rejected students
- Accepted students
- Institution
- Society

How do we design a machine learning algorithm that can optimize a complex combination of goals?

# Overall Goals for this Class

- Deeply explore bias and fairness research in ML.
  - Narrow focus.
  - Mostly original research papers.
- Practice being a successful academic.
  - Class targeted towards current PhD students with research interests in this area!

# Pre-Requisites

- This is not an introductory Machine Learning class!
- You need to satisfy one of the following ML pre-reqs:
  - Successfully taken CS4780 “Machine Learning”
  - Successfully taken CS6780 “Advanced Machine Learning”
  - Successfully taken a comparable “Intro to ML” class (\*)
  - Acquired the equivalent ML knowledge in some other way (e.g. strong background in Statistics + ML textbook) (\*)
- You need to be a PhD student
- Currently doing or planning to do research in this area of ML
- Basic probability, basic statistics, general mathematical maturity

(\*) means talk to me



# Format of Class

- Lectures (by TJ)
  - Background material
- Research paper presentations (by students)
  - Explore current state of the art
- Peer reviewing

# Research Paper Presentations

- Students present the paper in class
  - Slide presentation
  - Prepare discussion topics / group activity
  - Create critique, extended bibliography, examples, demo software, experiments etc. that help understand the paper
  - Prepare quiz
- Everybody reads the paper in preparation for class
  - Quiz about each paper
- All students give feedback afterwards.

# Peer Reviewing

- Goals
  - Give presenter constructive feedback from audience.
  - Reviewer has to think through what works about a presentation.
  - Learn how to write reviews. Be constructive, respectful, and mindful of biases.
- Reviewing the reviewers
  - Presenter gets to give feedback on the reviews (both direct and confidential to me)

# Credit Options and Grades

- Pass/Fail: Need to get at least 50% of points on each of following to pass.
  - paper presentation
  - in-class quizzes (lowest grades replaced by second lowest grade)
  - peer reviewing (lowest grades replaced by second lowest grade)
  - in-class participation
- Letter grade:
  - not allowed
- Audit:
  - not allowed, unless you have very good arguments

# Course Material

- Reference Books
  - Imbens, Rubin, "Causal Inference for Statistics, Social, and Biomedical Sciences", Cambridge University Press, 2015. ([online](#) via Cornell Library)
  - Morgan, Winship "Counterfactuals and Causal Inference", Cambridge University Press, 2007.
  - Barocas, Hardt, Narayanan. "Fairness and Machine Learning". ([online](#))
- Background Reading
  - K. Murphy, "Machine Learning - a Probabilistic Perspective", MIT Press, 2012. ([online](#) via Cornell Library)
  - Shai Shalev-Shwartz, Shai Ben-David, "Understanding Machine Learning - From Theory to Algorithms", Cambridge University Press, 2014. ([online](#))
  - B. Schoelkopf, A. Smola, "Learning with Kernels", MIT Press, 2001. ([online](#))
  - C. Bishop, "Pattern Recognition and Machine Learning", Springer, 2006.
  - R. Duda, P. Hart, D. Stork, "Pattern Classification", Wiley, 2001.
  - T. Hastie, R. Tibshirani, and J. Friedman, "The Elements of Statistical Learning", Springer, 2001.
  - T. Joachims, A. Swaminathan. SIGIR Tutorial on Counterfactual Evaluation and Learning for Search, Recommendation and Ad Placement, 2016. ([homepage](#))
- Slides, Notes and Papers
  - Slides available on course homepage or CMT
  - Papers on course homepage

# Bidding on Dates to Present

- Use CMT bidding mechanism to reserve dates to presenters
  - If you are
    - enrolled via studentcenter,
    - filled out the paper sheet (no promise we still have space though)you will get email from me through CMT.
  - Place your bids on the dates by Friday night.
  - I'll send you your assignment next week.
- Two weeks before your date
  - You get in touch with me and we finalize what you will be presenting.
  - I'll have suggestions for papers, but I am happy to follow your suggestions as well.

# How to Get in Touch

- Course Web Page
  - <https://www.cs.cornell.edu/Courses/cs7792/2020sp/>
- Email
  - Thorsten Joachims: [tj@cs.cornell.edu](mailto:tj@cs.cornell.edu)
- Office Hours
  - Wednesdays 11:10pm – 12:00pm, 418 Gates Hall
- Piazza
  - <https://piazza.com/cornell/spring2020/cs7792>
- Peer reviewing platform
  - <https://cmt3.research.microsoft.com/CS77922020>