Data Mining: Final Topics

CS4821-CS5831-s24

Laura E. Brown

In this class …

- You studied algorithms that exploit and reveal patterns in data.
- Goal:
  - You learned how to think about problems in data mining
  - You learned about a set of data analysis tools:
    - How to use them
    - What their assumptions are
    - The capabilities and limitations

Methods Examined

- Exploratory Data Analysis and Visualization
- Supervised learning
    - Classification
- Unsupervised learning
    - Clustering
    - Data Reduction
- Association Mining
- Recommendation Systems

Lessons Learned

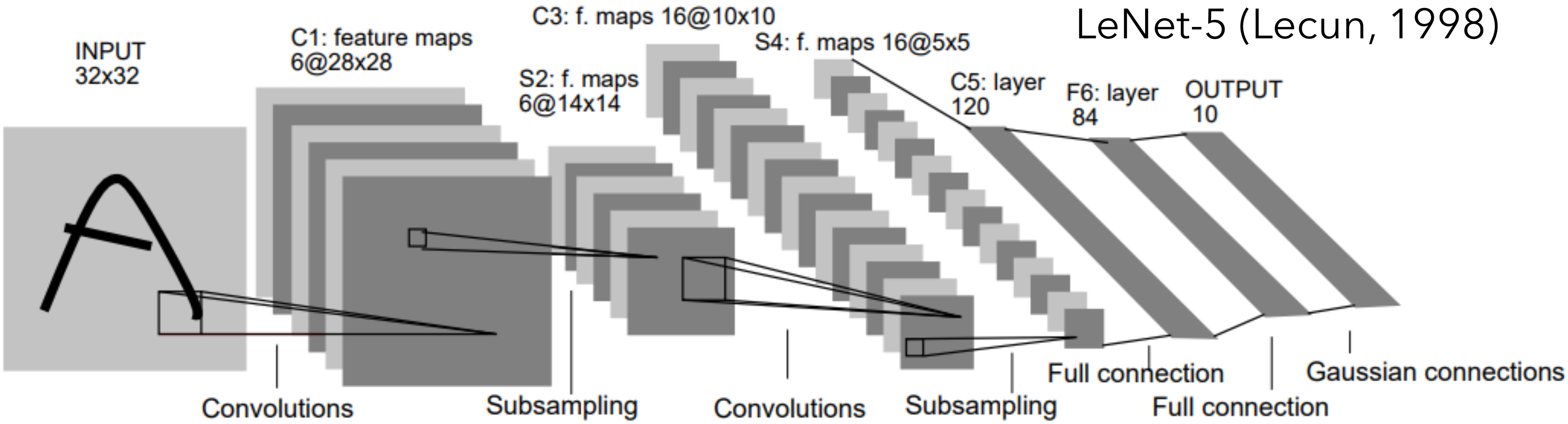
- Follow best practice in data preparation, testing and evaluation
- Use standard libraries/packages once data is clean and in proper form
  - Getting data prepared is time consuming
- Prediction results depend on data / problem
- Think about all aspects of how data is being collected and used

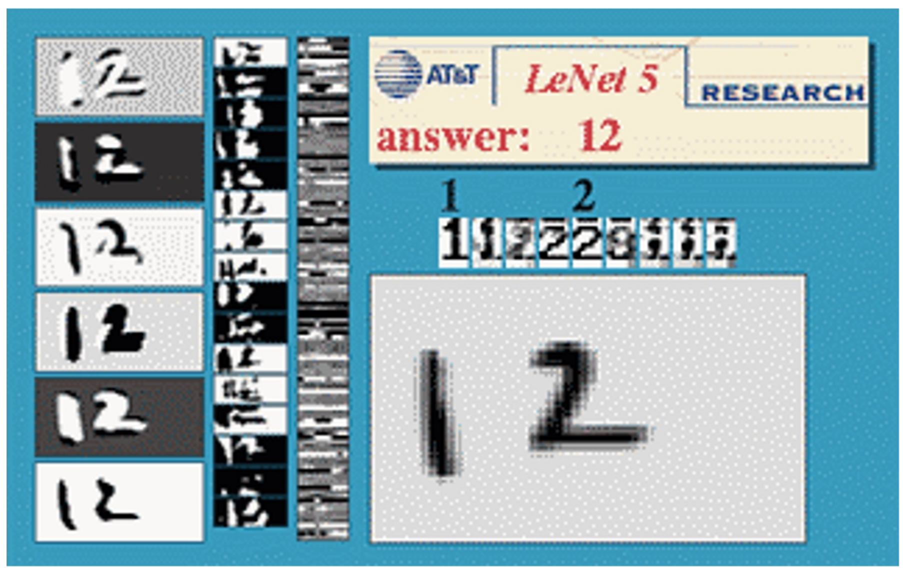
Challenges of Data Mining (ML)

- Opaqueness / Explainability / Interpretability
- Transparency / Privacy
- Algorithmic Bias

Opaqueness / Explainability

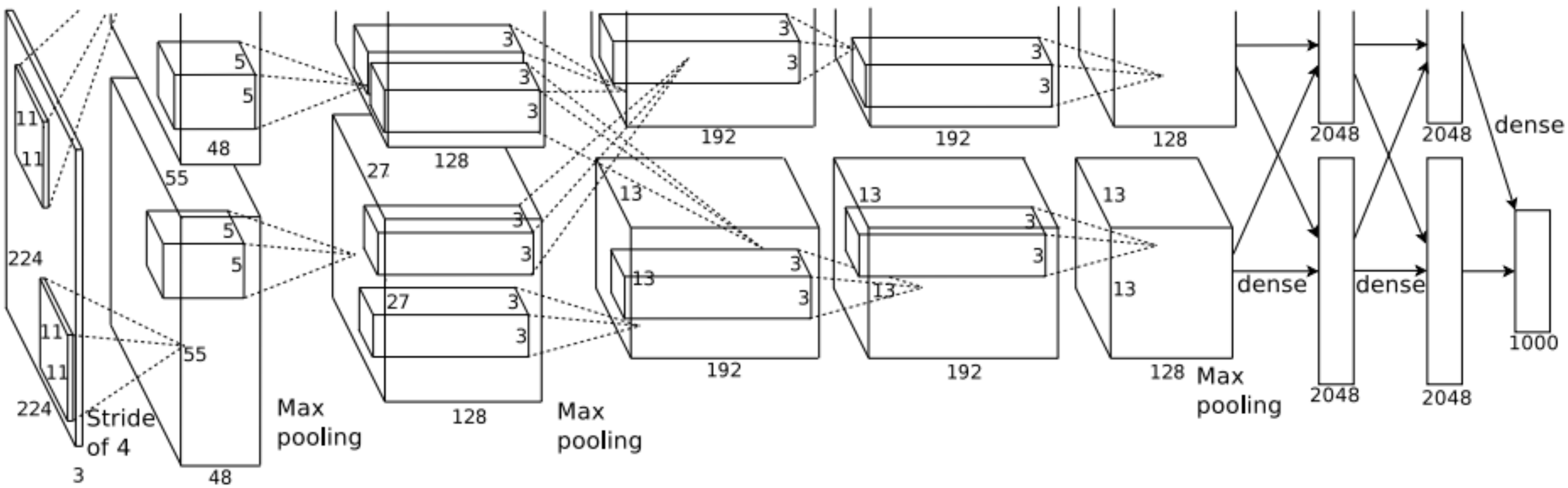
LeNet-5 (Lecun, 1998)



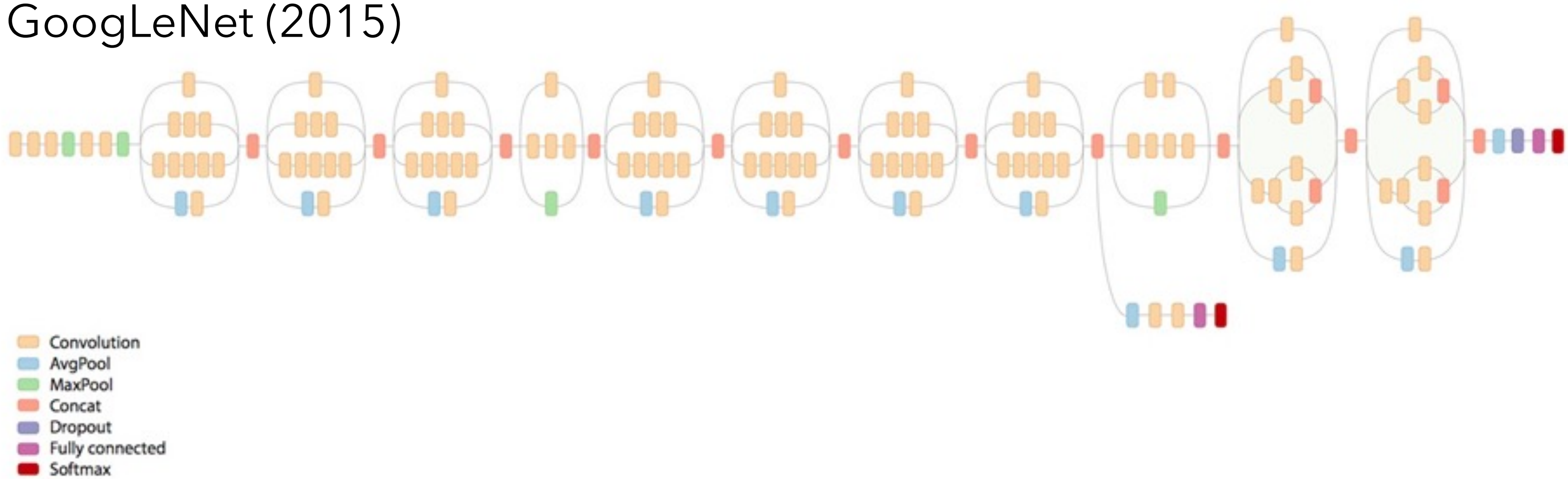


Opaqueness / Explainability

AlexNet (Hinton, 2012)



GoogLeNet (2015)



Another view of GoogLeNet's architecture.

Opaqueness / Explainability

Being able to explain

why or how a prediction is made

can be as or more important

than high predictive performance

Challenges of DM / ML

- **Transparency**
  - We often don't know the full amount of information about us that is available for companies/governments to use
  - Moreover, what information may be used to create a ML model is often proprietary and won't be disclosed even under FOIA (Freedom of Information Act)

Transparency

- Can information collected for one purpose be used for data in another purpose?
  - In Europe, generally no, not without explicit consent
  - In US, generally yes
- Companies routinely collect information about customers / employees and use it for marketing, sell it, …

Examples

WSJ, 2016

- Castlight Health, and other companies, offer employers access to information about their employees
  - Data in aggregate
- Launched a new product to predict pregnancy
- This information can flow from Castlight to employer

Examples

- How Companies Learn You Secrets
  NYTimes Magazine Feb. 16, 2012
  - Marketers want to send specialized ads to pregnant women
  - Target identified 25 products to predict pregnancy

Companies Tracking You

- Online
  - Cookies - WashingtonPost, June 21, 2019
- Phones
  - Apps, location data
- Banking / Purchases
- Smart TVs, Smart …

Challenges of Modern DM/ML

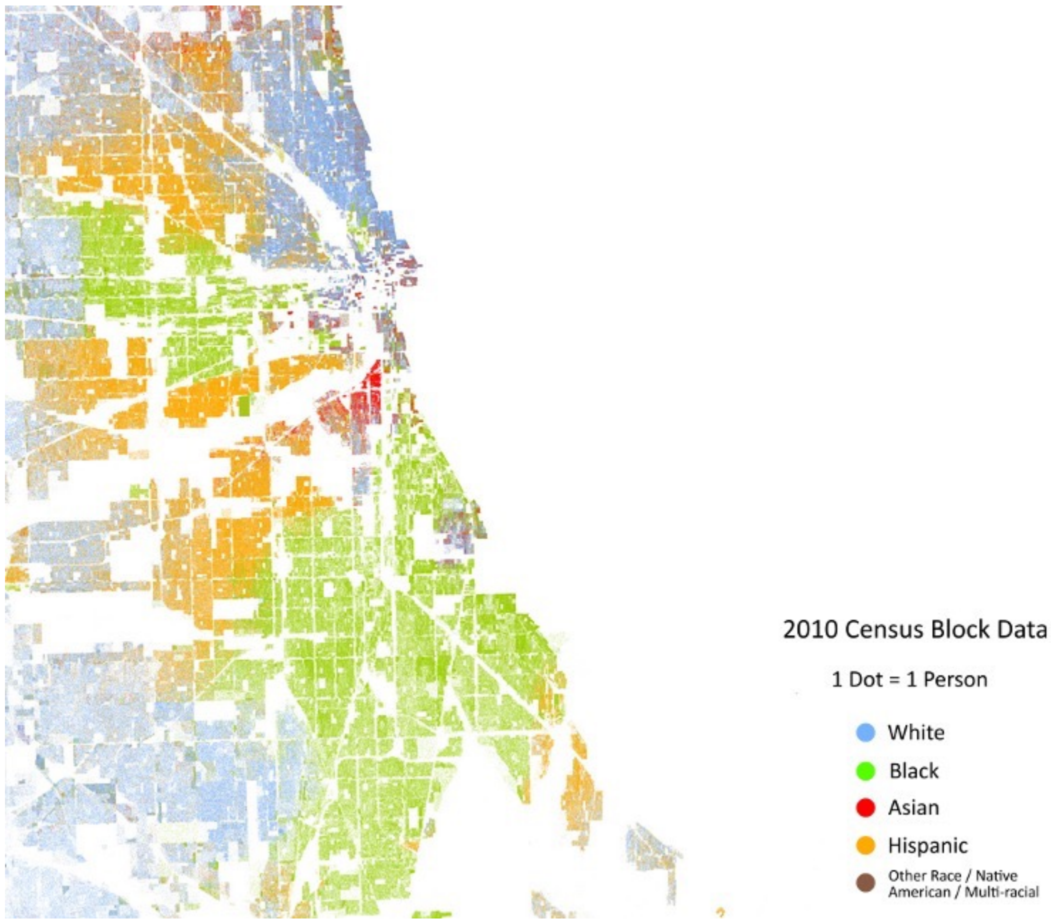
- **Algorithmic Bias**
- Example: You are creating a model for a bank to decide whether someone gets a mortgage
  - You know certain attributes are forbidden under the ECOA and FHA, e.g.,
    - Race
    - Religion
    - Sex
    - Age
    - …

Algorithmic Bias

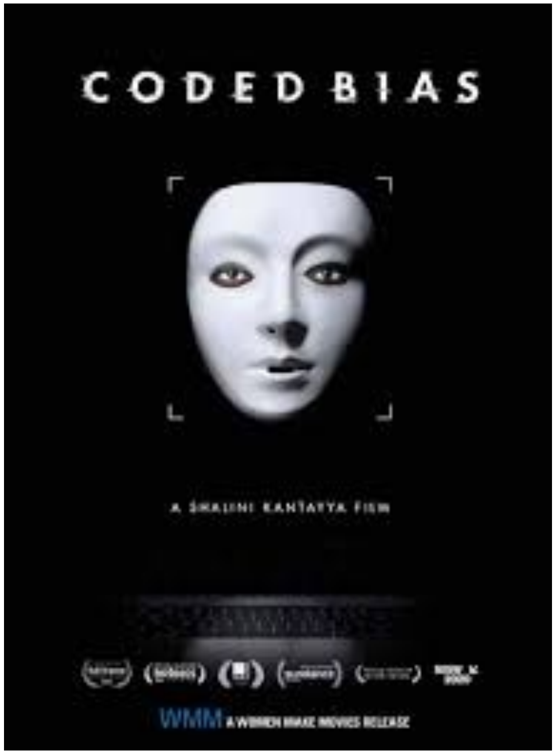
- Example: You are creating a model for a bank to decide whether someone gets a mortgage
  - You decide not to include these factors in your model
  - What is the problem?

Algorithmic Bias

- Other information correlated with these factors, e.g., race and zipcode



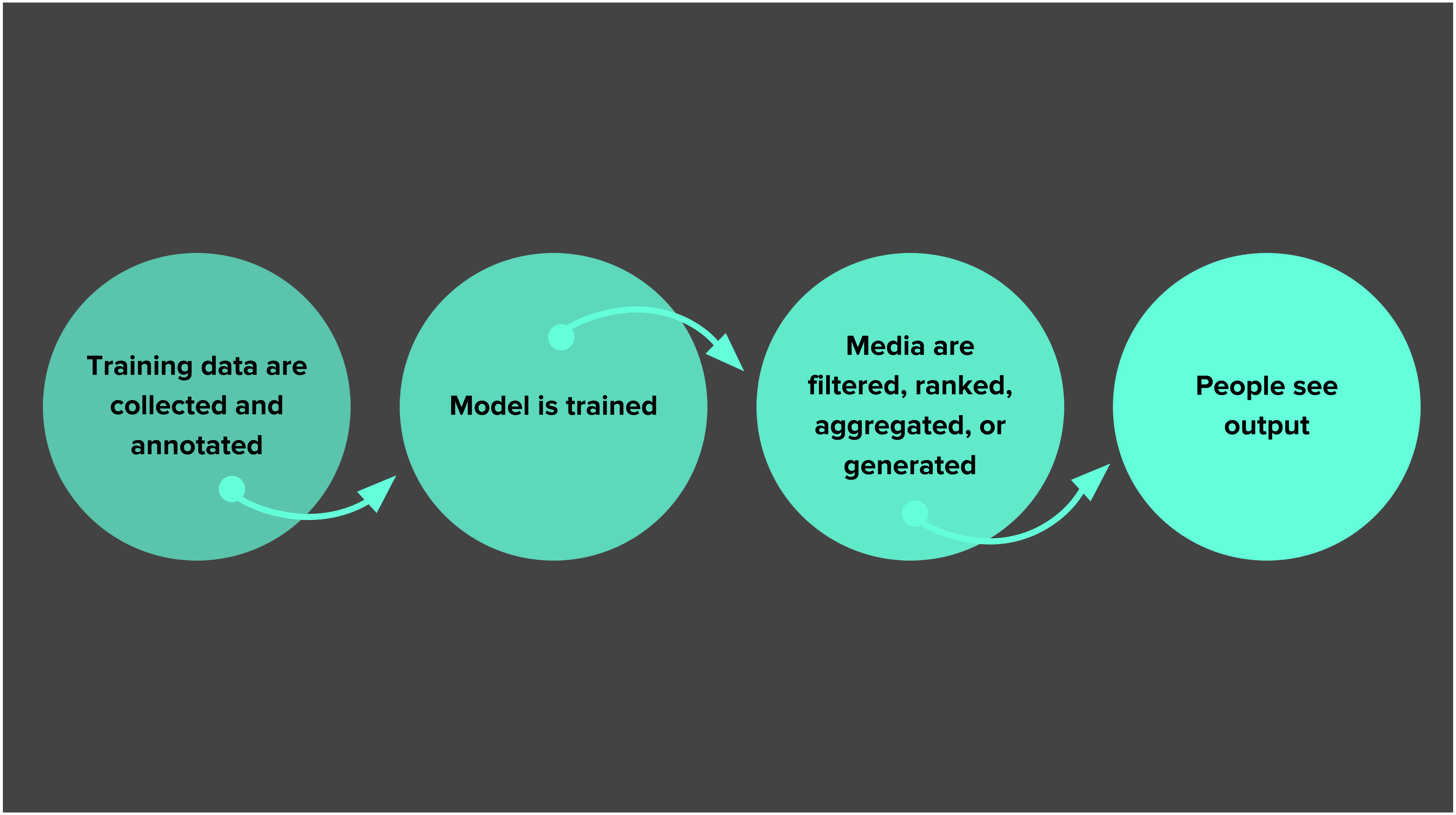
Coded Bias



Documentary on issues of bias in ML and algorithmic systems

Currently available on Netflix, PBS Passport

Location of Bias



Slide from M. Mitchell et al.

Location of Bias

Training data are collected and annotated

Human Biases in Data

- Reporting bias
- Selection bias
- Overgeneralization
- Out-group homogeneity bias
- Stereotypical bias
- Historical unfairness
- Implicit associations
- Implicit stereotypes
- Prejudice
- Group attribution error
- Halo effect

Human Biases in Collection and Annotation

- Sampling error
- Non-sampling error
- Insensitivity to sample size
- Correspondence bias
- In-group bias
- Bias blind spot
- Confirmation bias
- Subjective validation
- Experimenter's bias
- Choice-supportive bias
- Neglect of probability
- Anecdotal fallacy
- Illusion of validity

Slide from M. Mitchell et al.

Biases in Data

- **Selection Bias**: selection does not reflect a random sample

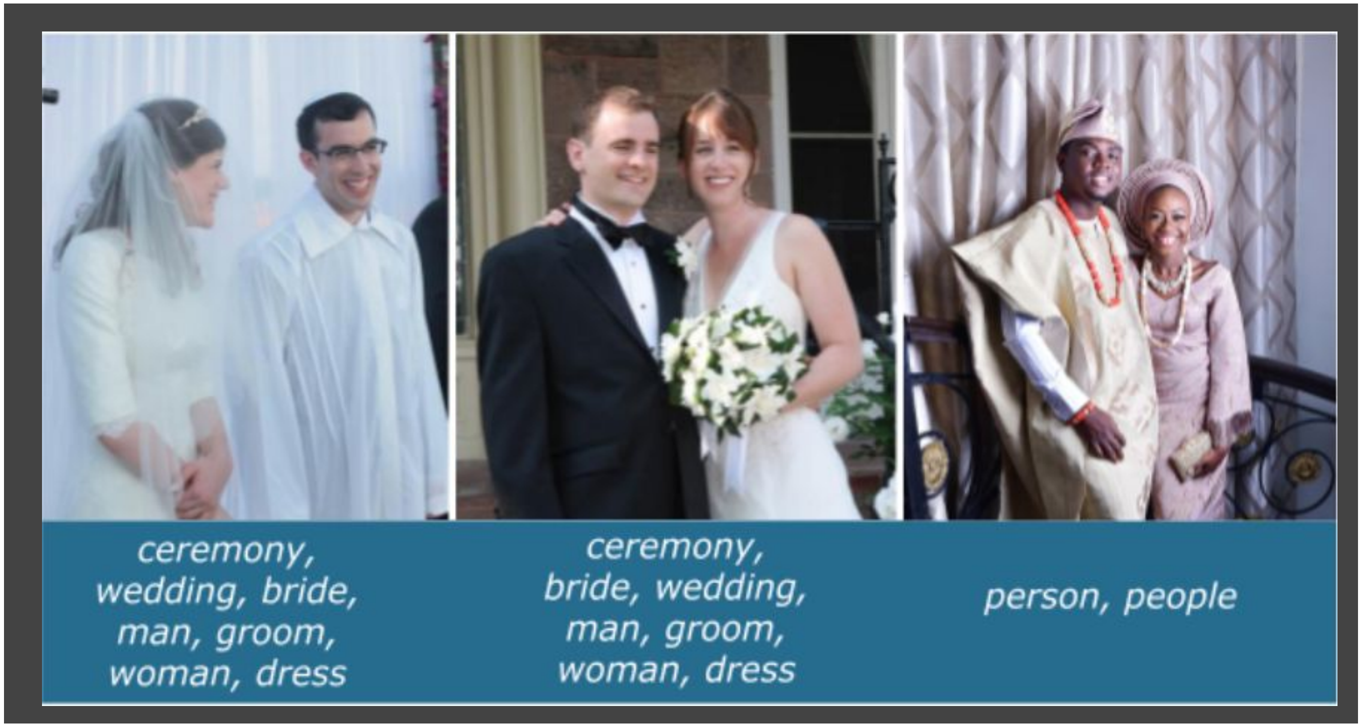
CREDIT

© 2013–2016 Michael Yoshitaka Erlewine and Hadas Kotek

Slide from M. Mitchell et al.

Biases in Data

- **Biased Labels** - annotations in data set will reflect worldviews of annotations

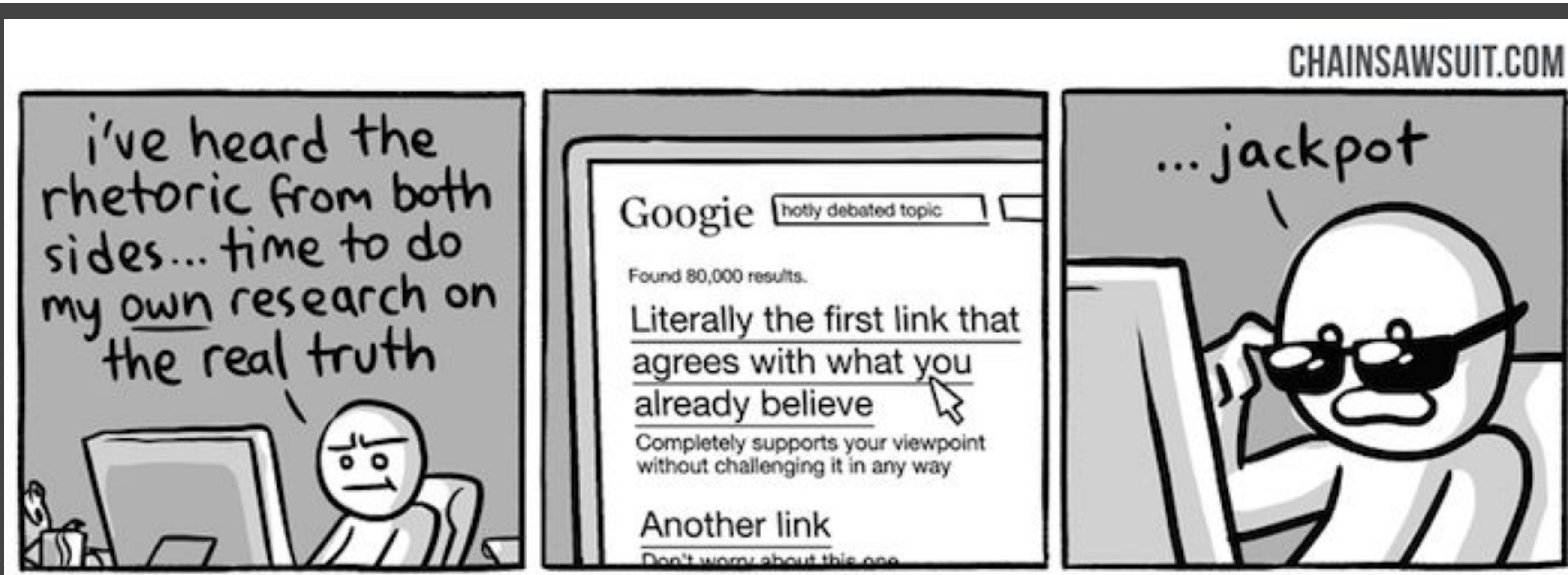


https://ai.googleblog.com/2018/09/introducing-inclusive-images-competition.html

Slide from M. Mitchell et al.

Biases in Interpretation

- **Confirmation bias**: The tendency to search for, interpret, favor, recall information in a way that confirms preexisting beliefs

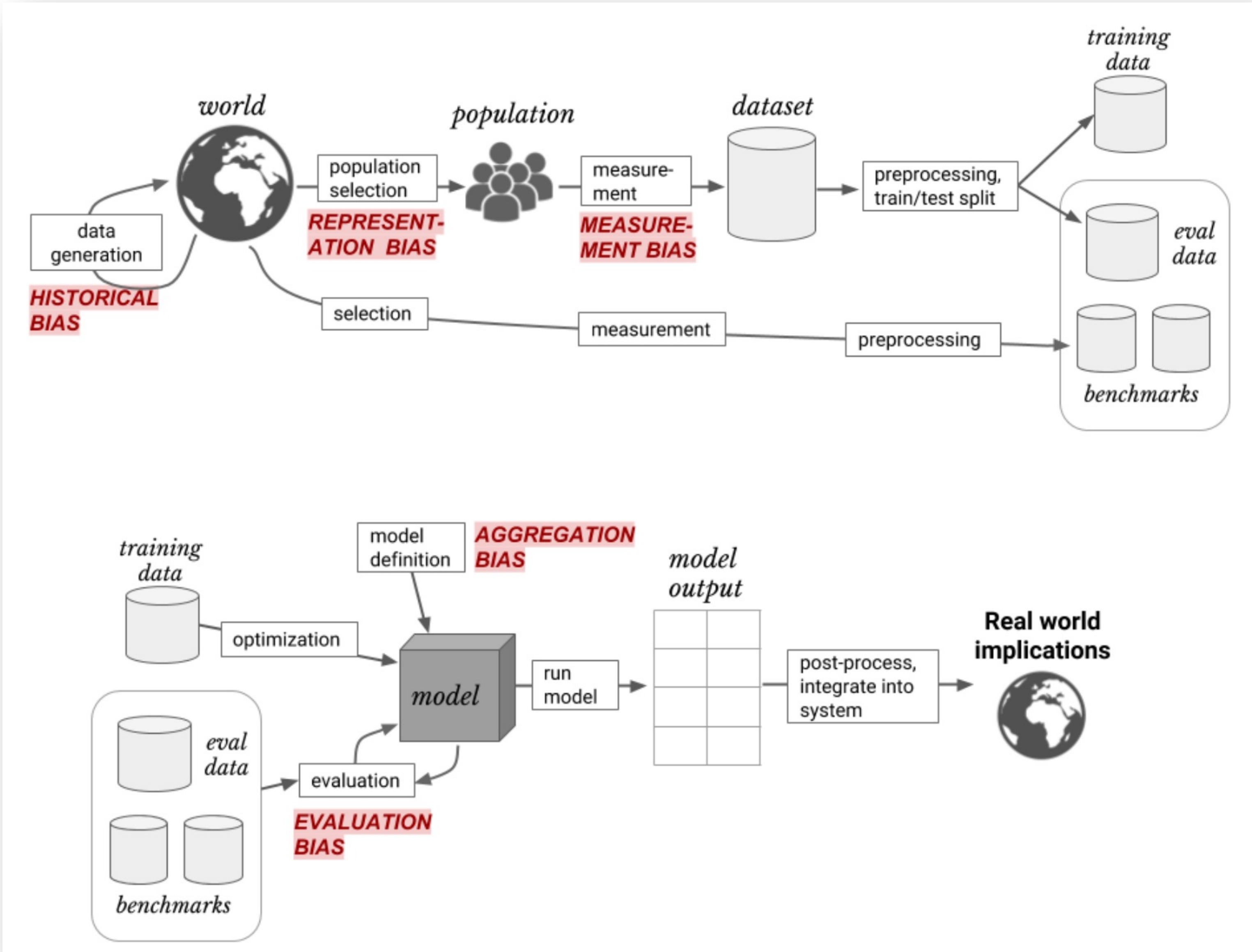


CREDIT
© kris straub - Chainsawsuit.com

Slide from M. Mitchell et al.

Another perspective

Different sources of bias have different causes

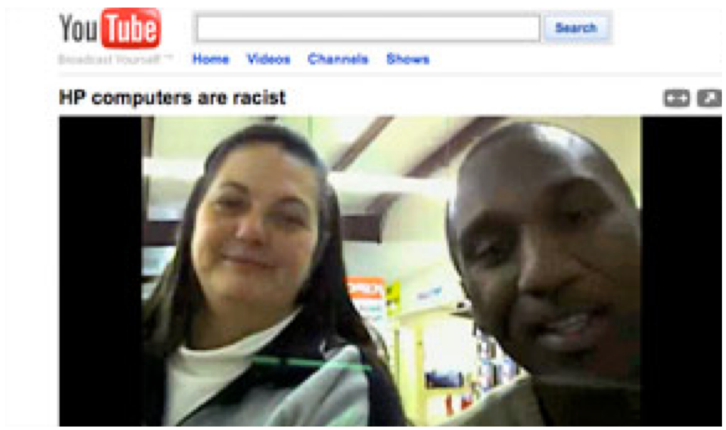


H. Suresh, J. Guttag

Algorithmic Bias

- Unjust, unfair, or prejudicial treatment of people related to race, income, sexual orientation, religion, gender, and other characteristics historically associated with discrimination and marginalization, when and where they manifest in algorithmic systems or algorithmically aided decision-making

Case study: Facial Recognition



HP Face-Tracking Webcams Don't Recognize Black People

Adam Frucci
12/21/09 10:00am • Filed to: WEBCAMS

154.6K 291

Slide from J. Kroll

Slide from J. Kroll

Gender Shades

Gender Classifier	Overall Accuracy on all Subjects in Pilot Parlaiments Benchmark (2017)
Microsoft	93.7%
FACE++	90.0%
IBM	87.9%

Gender Classifier	Female Subjects Accuracy	Male Subjects Accuracy	Error Rate Diff.
Microsoft	89.3%	97.4%	8.1%
FACE++	78.7%	99.3%	20.6%
IBM	79.7%	94.4%	14.7%

Joy Buolamwini & Timnit Gebru, gendershades.org

Slide from J. Kroll

Gender Shades

Gender Classifier	Darker Subjects Accuracy	Lighter Subjects Accuracy	Error Rate Diff.
Microsoft	87.1%	99.3%	12.2%
FACE++	83.5%	95.3%	11.8%
IBM	77.6%	96.8%	19.2%

Gender Classifier	Darker Male	Darker Female	Lighter Male	Lighter Female	Largest Gap
Microsoft	94.0%	79.2%	100%	98.3%	20.8%
FACE++	99.3%	65.5%	99.2%	94.0%	33.8%
IBM	88.0%	65.3%	99.7%	92.9%	34.4%

Joy Buolamwini & Timnit Gebru, gendershades.org

Slide from J. Kroll

Amazon Is Pushing Facial Technology That a Study Says Could Be Biased

In new tests, Amazon's system had more difficulty identifying the gender of female and darker-skinned faces than similar services from IBM and Microsoft.

Amazon's facial recognition matched 28 members of Congress to criminal mugshots

With No Laws To Guide It, Here's How Orlando Is Using Amazon's Facial Recognition Technology

Runaway Feedback Loops in Predictive Policing

Danielle Ensign, Sorelle A. Friedler, Scott Neville, Carlos Scheidegger, Suresh Venkatasubramanian

Slide from R. Thomas

Case study: Online Ad Delivery

Racism is Poisoning Online Ad Delivery, Says Harvard Professor

Facebook's ad system seems to discriminate by race and gender

Discrimination in Online Ad Delivery

Latanya Sweeney

Discrimination through optimization: How Facebook's ad delivery can lead to skewed outcomes

Muhammad Ali, Piotr Sapiezynski, Miranda Bogen, Aleksandra Korolova, Alan Mislove, Aaron Rieke

Facebook (Still) Letting Housing Advertisers Exclude Users by Race

Dozens of Companies Are Using Facebook to Exclude Older Workers From Job Ads

Ad related to **latanya sweeney**

Latanya Sweeney Truth
www.instantcheckmate.com/
Looking for **Latanya Sweeney**? Check Latar

Ads by Google

Latanya Sweeney, Arrested?
1) Enter Name and State. 2) Access Fu
Checks Instantly.
www.instantcheckmate.com/

Latanya Sweeney
Public Records Found For: **Latanya Sv**
www.publicrecords.com/

La Tanya
Search for La Tanya Look Up Fast Res
www.ask.com/La+Tanya

Ads by Google

Kirsten Lindquist
Get **Kirsten Lindquist** Find **Kirsten Lindquist**
www.ask.com/**Kirsten+Lindquist**

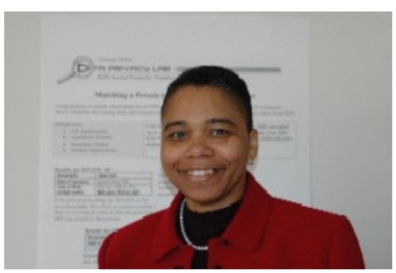
We Found:**Kristen Lindquist**
1) Contact **Kristen Lindquist** - Free Info! 2) Cu
Phone, Address & More.
www.peoplesmart.com/

Search by Phone | Search by Email
Background Checks | Search by Address
Public Records | Criminal Records

Kristen Lindquist
Public Records Found For: **Kristen Lindquist**. View Now.
www.publicrecords.com/

Racism is Poisoning Online Ad Delivery, Says Harvard Professor

Latanya Sweeney, Ph.D.



Facebook's ad system seems to discriminate by race and gender

New research shows that Facebook's ad-distribution software is disturbingly biased

L. Sweeney (2013). Discrimination in online ad delivery. Queue, 11(3). See also N. Newman (2011) in Huffington Post.

Slide from R. Thomas

Facebook (Still) Letting Housing Advertisers Exclude Users by Race

After ProPublica revealed last year that Facebook advertisers could target housing ads to whites only, the company announced it had built a system to spot and reject discriminatory ads. We retested and found major omissions.

by **Julia Angwin**, **Ariana Tobin** and **Madeleine Varner**, Nov. 21, 2017, 1:23 p.m. EST

Dozens of Companies Are Using Facebook to Exclude Older Workers From Job Ads

Among the companies we found doing it: Amazon, Verizon, UPS and Facebook itself. “It’s blatantly unlawful,” said one employment law expert.

Amazon reportedly scraps internal AI recruiting tool that was biased against women

The secret program penalized applications that contained the word “women’s”

By James Vincent | Oct 10, 2018, 7:09am EDT

Slide from R. Thomas

Gender Stereotypes

Google image search for "**C.E.O.**" produced 11 percent women, even though 27 percent of United States chief executives are women.

M. Kay, C. Matuszek, S. Munson (2015): [Unequal Representation and Gender Stereotypes in Image Search Results for Occupations](). CHI'15.

Slide from R. Thomas

Recent Examples

Problems

- Facebook Ad System Perpetuates Gender Bias, The Intercept April 2021, Research Site
- Proctorio Facial Recognition, Vice April 2021
- Automatic Speech Recognition Systems Exhibit Bias, VentureBeat April 2021

What about LLM System?

- Stable Diffusions text-to-image model amplifies stereotypes – Bloomberg, 2023
  https://www.bloomberg.com/graphics/2023-generative-ai-bias/

- OpenAI's GPT shows bias in resume ranking experiment – Bloomberg, 2024
  https://www.bloomberg.com/graphics/2024-openai-gpt-hiring-racial-discrimination/

Algorithms are used differently than human decision makers:

- Algorithms are often used **at scale**
- Algorithmic systems are **cheap**
- Algorithms are more likely to be implemented with **no appeals process** in place
- People are more likely to assume algorithms are **objective or error-free** (even if they're given the option of a human override)

The privileged are processed by people; the poor are processed by algorithms. (Cathy O'Neil)

Questions to Ask

- Should we even be doing this?
- What bias is in the data?
- Can the code and data be audited?
- What are error rates for different sub-groups?
- What is the accuracy of a simple rule-based alternative?
- What processes are in place to handle appeals or mistakes?
- How diverse is the team that built it?

Conclusion

- Bad news: the algorithms and big data are not just **mirroring** the existing bias but also they are **reinforcing** that bias and amplifying **inequality**

- Good news: there are still opportunities to build tools to address different aspects of this problem

Resources

- Conferences/Communities:
  - Fairness, Accountability, and Transparency in Machine Learning (FATML)
    - https://fatml.org and https://fatconference.org
  - Data Transparency Lab
    - https://dtlconferences.org
- Presentations:
  - S. Hajian, F. Bonchi, and C. Castillo at KDD 2016
  - S. Venkatasubramanian at ICWSM 2016
  - K. Crawford at NIPS2017, https://www.youtube.com/watch?v=fMym_BKWQzk

Resources

- Cathy O' Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, 2016
- "Technology is Biased too. How do we Fix it?" By Laura Hudson (Five Thirty Eight, July 20, 2017).
- "Principles for Accountable Algorithms and a Social Impact Statement for Algorithms" by Nicholas Diakopoulos et al. (Fairness, Accountability, and Transparency in Machine Learning)