Data Science 100
Lecture 27: Ethics in Data Science

Slides by:
Joshua Kroll
kroll@berkeley.edu
HP Face-Tracking Webcams Don't Recognize Black People

Adam Frucci
12/21/09 10:00am • Filed to: WEBCAMS
'A white mask worked better': why algorithms are not colour blind

When Joy Buolamwini found that a robot recognised her face better when she wore a white mask, she knew a problem needed fixing.
## Gender Shades

<table>
<thead>
<tr>
<th>Gender Classifier</th>
<th>Overall Accuracy on all Subjects in Pilot Parliaments Benchmark (2017)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Microsoft</strong></td>
<td>93.7%</td>
</tr>
<tr>
<td><strong>FACE</strong></td>
<td>90.0%</td>
</tr>
<tr>
<td><strong>IBM</strong></td>
<td>87.9%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gender Classifier</th>
<th>Female Subjects Accuracy</th>
<th>Male Subjects Accuracy</th>
<th>Error Rate Diff.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Microsoft</strong></td>
<td>89.3%</td>
<td>97.4%</td>
<td>8.1%</td>
</tr>
<tr>
<td><strong>FACE</strong></td>
<td>78.7%</td>
<td>99.3%</td>
<td>20.6%</td>
</tr>
<tr>
<td><strong>IBM</strong></td>
<td>79.7%</td>
<td>94.4%</td>
<td>14.7%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gender Classifier</th>
<th>Darker Subjects Accuracy</th>
<th>Lighter Subjects Accuracy</th>
<th>Error Rate Diff.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Microsoft</strong></td>
<td>87.1%</td>
<td>99.3%</td>
<td>12.2%</td>
</tr>
<tr>
<td><strong>FACE</strong></td>
<td>83.5%</td>
<td>95.3%</td>
<td>11.8%</td>
</tr>
<tr>
<td><strong>IBM</strong></td>
<td>77.6%</td>
<td>96.6%</td>
<td>19.2%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gender Classifier</th>
<th>Darker Male</th>
<th>Darker Female</th>
<th>Lighter Male</th>
<th>Lighter Female</th>
<th>Largest Gap</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Microsoft</strong></td>
<td>94.0%</td>
<td>79.2%</td>
<td>100%</td>
<td>98.3%</td>
<td>20.8%</td>
</tr>
<tr>
<td><strong>FACE</strong></td>
<td>99.3%</td>
<td>65.5%</td>
<td>99.2%</td>
<td>94.0%</td>
<td>33.8%</td>
</tr>
<tr>
<td><strong>IBM</strong></td>
<td>88.0%</td>
<td>65.3%</td>
<td>99.7%</td>
<td>92.9%</td>
<td>34.4%</td>
</tr>
</tbody>
</table>
Google Photos Mistakenly Labels Black People ‘Gorillas’

BY CONOR DOUGHERTY  JULY 1, 2015 7:01 PM  📰 41
Data science is everywhere...

Scoring Systems
FICO
ChexSystems

Lotteries

Criminal Justice & National Security
TSA Pre✓

News

Lending
zest
earnest
Kabbage

Hiring
gild
TalentBin

NarrativeScience
Affirm
entelo

GJ
Data Science Life Cycle

Context
Question
Refine Question to an one answerable with data

Model evaluation
Prediction error

Design
Data Collection
Data Cleaning

Modeling
Test-train split
Loss function choice
Feature engineering
  Transformations, Dummy Variables
Model selection
  Best subset regression
  Cross-Validation

BIAS!
Extreme Digital Vetting of Visitors to the U.S. Moves Forward Under a New Name

ICE officials have invited tech companies, including Microsoft, to develop algorithms that will track visa holders’ social media activity.

by George Joseph, Nov. 22, 8 a.m. EST
A Stanford scientist says he built a gaydar using “the lamest” AI to prove a point
Proxies
Automated Inference on Criminality using Face Images

Xiaolin Wu
McMaster University
Shanghai Jiao Tong University
xwu510@gmail.com

Xi Zhang
Shanghai Jiao Tong University
zhangxi.19930818@sjtu.edu.cn

High IQ  Academic Researcher  Professional Poker Player  Terrorist
Data Collection

Street Bump

Help keep your streets smooth

RECORD A TRIP

MY TRIPS

Trip Summary

10 minutes ago

MILES - 1.1

13 BUMPS
Data Collection
Data Cleaning
Data Cleaning
Data Cleaning
Test-Train Split
DOES UNCONSCIOUS RACIAL BIAS AFFECT TRIAL JUDGES?

Jeffrey J. Rachlinski,* Sheri Lynn Johnson,† Andrew J. Wistrich,‡ & Chris Guthrie††
Machine Bias

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

by Julie Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica
May 23, 2016

On a spring afternoon in 2014, Brisha Borden was running late to pick up her god-sister from school when she spotted an unlocked kid's blue Huffy bicycle and a silver Razor scooter. Borden and a friend grabbed the bike and scooter and tried to ride them down the street in the Fort Lauderdale suburb of Coral Springs.
Hypothetical conditional effect plot, biased against blacks

Probability of two-year recidivism vs. COMPAS score decile

- White
- Black
So why does ProPublica think COMPAS is biased?
Feature Engineering
Feature Engineering
Feature Engineering
Model Selection
Model Evaluation

Age of Miss America correlates with Murders by steam, hot vapours and hot objects.
Model Evaluation

Letters in Winning Word of Scripps National Spelling Bee correlates with Number of people killed by venomous spiders

- For 15-letter words:
  - 1999: 15 deaths
  - 2000: 15 deaths
  - 2001: 15 deaths
  - 2002: 15 deaths
  - 2003: 15 deaths
  - 2004: 15 deaths
  - 2005: 15 deaths
  - 2006: 15 deaths
  - 2007: 15 deaths
  - 2008: 15 deaths
  - 2009: 15 deaths

- For 10-letter words:
  - 2000: 15 deaths
  - 2001: 15 deaths
  - 2002: 15 deaths
  - 2003: 15 deaths
  - 2004: 15 deaths
  - 2005: 15 deaths
  - 2006: 15 deaths
  - 2007: 15 deaths
  - 2008: 15 deaths
  - 2009: 15 deaths

- For 5-letter words:
  - 2000: 15 deaths
  - 2001: 15 deaths
  - 2002: 15 deaths
  - 2003: 15 deaths
  - 2004: 15 deaths
  - 2005: 15 deaths
  - 2006: 15 deaths
  - 2007: 15 deaths
  - 2008: 15 deaths
  - 2009: 15 deaths

Legend:
- Black line: Number of people killed by venomous spiders
- Red line: Spelling Bee winning word
So what do we do about this?
Ethical Lessons

1. **Know what the data are telling you**
   - Understand your model, and validate it
   - Involve domain experts; “raw data” is an oxymoron
   - Communicate not just model *capabilities*, but also *limits* and *assumptions*

2. **Data about people are about people**
   - The data do not exist on their own, someone made them

3. **Always** consider why your model might be wrong in systematic ways
   - Don’t just **trust** the technology, make it **trustworthy**
   - **Close** the data science lifecycle: the world changes, and so should your models
Ethics beyond the Data Science Lifecycle
Should we deploy technology at all?
When is it ethical to collect data?
When is it ethical to collect sensitive data?

• Without knowing sensitive attributes, it’s not possible to know how to make good decisions.
  • E.g., in a credit-granting context, there’s no reason to believe that qualified minority applicants will be similar to qualified members of the majority

• But having sensitive data means it can be repurposed for other uses
Inclusion and Representation
Resources

• Take a course at Berkeley on people and technology
• Visit https://fatml.org and https://fatconference.org
  • Reading list
  • Principles for Accountable Algorithms
• Books
  • Weapons of Math Destruction, by Cathy O’Neil
  • How to Lie With Statistics, by Darrell Huff
• Courses around the world: tinyurl.com/ethics-classes
“Remember that all models are wrong; the practical question is how wrong do they have to be to not be useful.”

-George E. P. Box