2020 PREDICTIVE ANALYTICS 4.0
VIRTUAL EVENT

JONATHAN CULBERT

S01 Session 1: Morality in the Machine: Ethics and the Rise of AI in the Insurance Industry

September 24, 2020

SOCIETY OF ACTUARIES®

SOCIETY OF ACTUARIES
Antitrust Compliance Guidelines

Active participation in the Society of Actuaries is an important aspect of membership. While the positive contributions of professional societies and associations are well-recognized and encouraged, association activities are vulnerable to close antitrust scrutiny. By their very nature, associations bring together industry competitors and other market participants.

The United States antitrust laws aim to protect consumers by preserving the free economy and prohibiting anti-competitive business practices; they promote competition. There are both state and federal antitrust laws, although state antitrust laws closely follow federal law. The Sherman Act, is the primary U.S. antitrust law pertaining to association activities. The Sherman Act prohibits every contract, combination or conspiracy that places an unreasonable restraint on trade. There are, however, some activities that are illegal under all circumstances, such as price fixing, market allocation and collusive bidding.

There is no safe harbor under the antitrust law for professional association activities. Therefore, association meeting participants should refrain from discussing any activity that could potentially be construed as having an anti-competitive effect. Discussions relating to product or service pricing, market allocations, membership restrictions, product standardization or other conditions on trade could arguably be perceived as a restraint on trade and may expose the SOA and its members to antitrust enforcement procedures.

While participating in all SOA in person meetings, webinars, teleconferences or side discussions, you should avoid discussing competitively sensitive information with competitors and follow these guidelines:

- Do not discuss prices for services or products or anything else that might affect prices
- Do not discuss what you or other entities plan to do in a particular geographic or product markets or with particular customers.
- Do not speak on behalf of the SOA or any of its committees unless specifically authorized to do so.
- Do leave a meeting where any anticompetitive pricing or market allocation discussion occurs.
- Do alert SOA staff and/or legal counsel to any concerning discussions
- Do consult with legal counsel before raising any matter or making a statement that may involve competitively sensitive information.

Adherence to these guidelines involves not only avoidance of antitrust violations, but avoidance of behavior which might be so construed. These guidelines only provide an overview of prohibited activities. SOA legal counsel reviews meeting agenda and materials as deemed appropriate and any discussion that departs from the formal agenda should be scrutinized carefully. Antitrust compliance is everyone's responsibility; however, please seek legal counsel if you have any questions or concerns.

Presentation Disclaimer

Presentations are intended for educational purposes only and do not replace independent professional judgment. Statements of fact and opinions expressed are those of the participants individually and, unless expressly stated to the contrary, are not the opinion or position of the Society of Actuaries, its cosponsors or its committees. The Society of Actuaries does not endorse or approve, and assumes no responsibility for, the content, accuracy or completeness of the information presented. Attendees should note that the sessions are audio-recorded and may be published in various media, including print, audio and video formats without further notice.

About me: broadcasting to actuarial science to data science

ENGINEERING

- Attended Univ. of Michigan; left school to pursue radio broadcasting career
- News anchor, reporter, and morning show host in Detroit
- Returned to school to finish degree at Eastern Michigan
- At Delta Dental of Michigan, Ohio, and Indiana since 2010
- Master of Science in Business Analytics at Carnegie Mellon University, expected graduation May 2021
  - Studying Machine Learning, Optimization, and data-driven leadership
  - Especially interested in questions of fairness and bias in machine learning

Carnegie Mellon University

2020 PREDICTIVE ANALYTICS 4.0 VIRTUAL EVENT

Foundations: what we mean when we say "Artificial Intelligence" and "Machine Learning"

SOCIETY OF ACTUARIES®

Are “AI” and “ML” just buzzwords?

Artificial Intelligence

Machine Learning

These terms are popular in marketing materials, but what do they *really* mean?

"Artificial Intelligence" is a slippery term

I like Andrew Moore's definition of Artificial Intelligence because it points out that the definition of "AI" is subjective and changes over time.

"Artificial intelligence is the science and engineering of making computers behave in ways that, until recently, we thought required human intelligence."

- Andrew Moore, Google Cloud (former dean of CMU School of Computer Science)

Quote from: P. High, "Carnegie Mellon Dean of Computer Science on the Future of AI," *Forbes*, 30 October 2017

What counts as “AI” has changed over time

ELIZA:

Primitive “chatbot” with simple logic

“Hey, Siri…”

AI built on deep machine learning

1965	1997	2011

Deep Blue:

“Rule based” system on a massive scale

Who writes the code for these AI programs?

"Hey Siri, find pictures of cats" + Magic? = Kitty!

Tasks like computer speech recognition improved significantly when technology moved from rules-based programs to AI built on machine learning.

Today’s AI is built on machine learning

Rules-Based

Advancements in AI

Machine Learning

1997

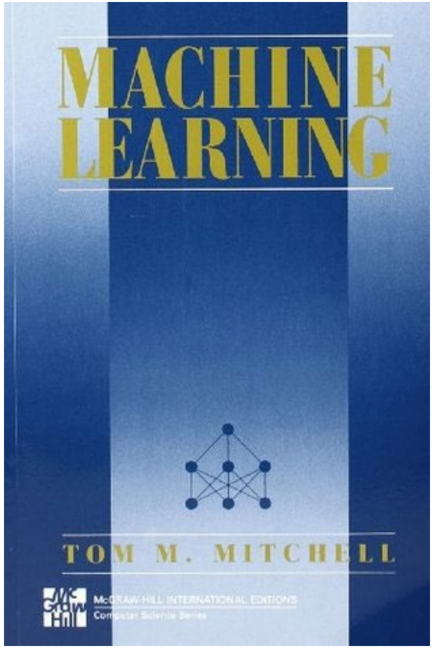
The leap from Deep Blue to AlphaGo represents a revolution in deep learning.

2016

Definition: In Machine Learning, programs "learn" from data

*"Machine learning is the study of computer algorithms that allow computer programs to **automatically improve** through **experience**."*

- Tom Mitchell, founding Chair of the Machine Learning Department, Carnegie Mellon University

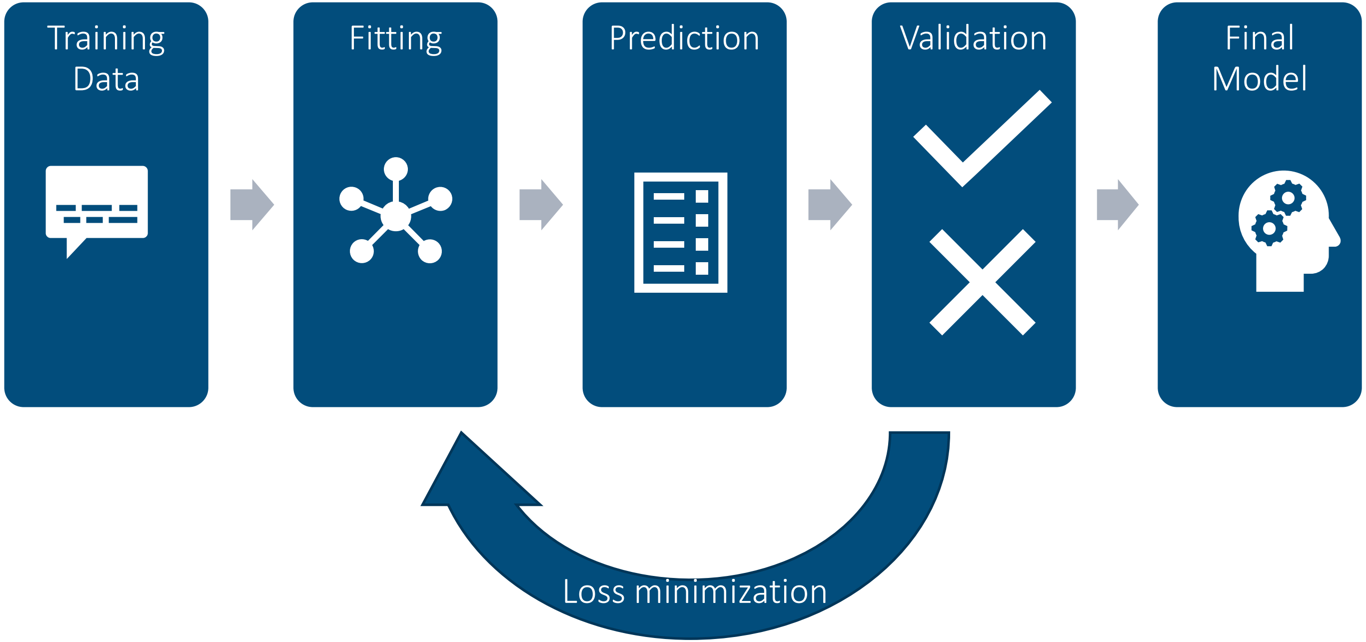


Quote and cover image from: T. Mitchell, *Machine Learning*, McGraw Hill, 1997.

(Thanks to Prof. Zachary C. Lipton of Carnegie Mellon University for recommending these definitions of AI and ML.)

(Supervised) machine learning is much like regression, with extremely expressive functional forms

A simplified view of supervised learning



Key idea:

To understand ethical concerns in ML and AI, remember that ML models are highly dependent on their training data.

2020 PREDICTIVE ANALYTICS 4.0 VIRTUAL EVENT

Applications of Machine Learning and the ethical questions they invite

SOCIETY OF ACTUARIES®

AI is all around us

Human Resources

npr

TECHNOLOGY

Can Artificial Intelligence Make The Hiring Process More Fair?

April 8, 2019 · 4:40 PM ET
Heard on All Things Considered

REUTERS

TECHNOLOGY 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

M. Savage, "Can Artificial Intelligence Make The Hiring Process More Fair?," *All Things Considered*, NPR, 8 April 2019.
J. Dastin, "Amazon scraps secret AI recruiting tool that showed bias against women," *Reuters*, 9 October 2018.

AI is all around us

Autonomous Vehicles

AI & ML in insurance

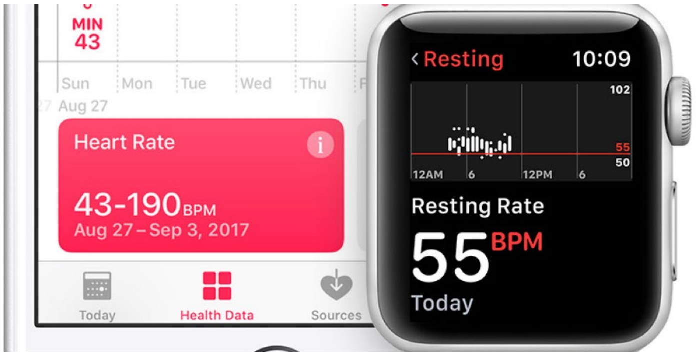
Auto Insurance

- Telematic Data
- Crash Damage Recognition

AI & ML in insurance

Health Insurance

- Biometric Data
- Fraud Detection



AI & ML in insurance

Across insurance lines

- Product Design
- Marketing
- Claims Processing

Hopes that computer algorithms are "automatically fair" are naïve

Moritz Hardt (UC Berkeley), leading ML fairness researcher:

Machine learning acts as a social mirror, reflecting and sometimes amplifying society's inequities.

M. Hardt, "How big data is unfair," *Medium*, 14 September 2014.

Key idea:

Algorithms aren't automatically fair. It is the difficult task of machine learning practitioners to identify and correct algorithmic bias.

Unintentional bias example:
Hiring AI favored men, resisting correction efforts

REUTERS

TECHNOLOGY 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

"They literally wanted it to be an engine where I'm going to give you 100 resumes, it will spit out the top five, and we'll hire those."

"(The algorithm) penalized resumes that included the word 'women's,' as in 'women's chess club captain.'"

J. Dastin, "Amazon scraps secret AI recruiting tool that showed bias against women," *Reuters*, 9 October 2018.

What happened?
Past imbalance led to unwanted bias in new model

The model was trained on historical hiring data, and past hiring had favored men.

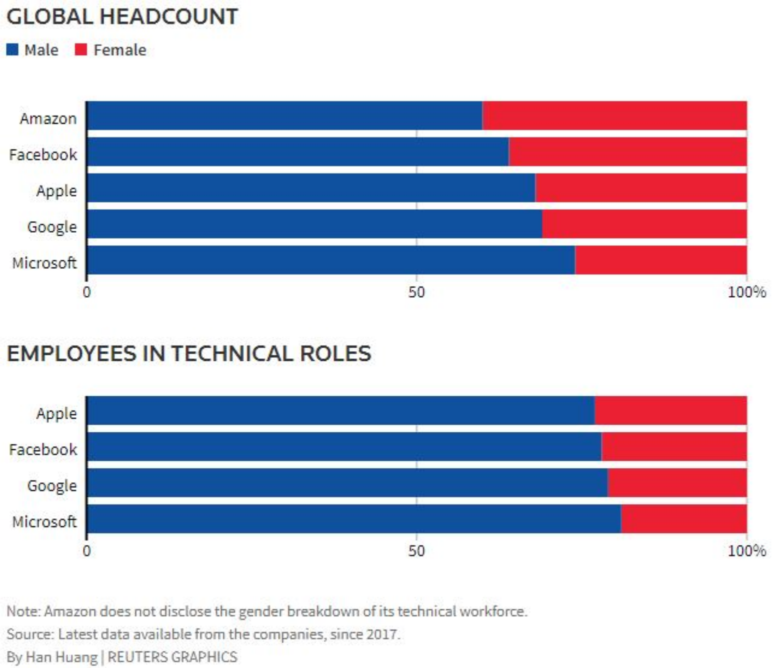


Image by Han Huang, Reuters Graphics

Bias persisted even as gender was hidden

Bias persisted even as gender was hidden

Traits like gender and race can leak into the model through correlated features

Known as "redundant encoding," sensitive attributes are encoded in the model via proxies or other relationships.

Actuaries will be familiar with strong proxies like 5-digit ZIP Code. There are many other, more subtle examples.

Principal: No fairness through unawareness [1]

Efforts to remove redundant encoding can be ineffective and even harmful to protected groups. [2]

[1] S. Barocas, M. Hardt and A. Narayanan, Fairness in Machine Learning, fairmlbook.org, 2019.
[2] Z. C. Lipton, A. Chouldechova and J. McAuley, "Does mitigating ML's impact disparity require treatment disparity?," in Advances in Neural Information Processing Systems, 2018.

Example: Program to control healthcare costs unintentionally included racial bias

RESEARCH ARTICLE

Science

ECONOMICS

Dissecting racial bias in an algorithm used to manage the health of populations

Ziad Obermeyer[1,2*], Brian Powers[3], Christine Vogeli[4], Sendhil Mullainathan[5*†]

Z. Obermeyer, et al. "Dissecting racial bias in an algorithm used to manage the health of populations." *Science* 366.6464 (2019): 447-453.

Example: Program to control healthcare costs unintentionally included racial bias

- Risk scores used for high-risk care management program
- Researchers found enhanced care was disproportionately offered to White patients compared to similarly sick Black patients
- The source of bias was subtle. We'll revisit this later on.

Z. Obermeyer, et al. "Dissecting racial bias in an algorithm used to manage the health of populations." *Science* 366.6464 (2019): 447-453.

Key idea:

Machine learning bias is rarely caused by malicious actors. It is almost always **unintentional** and therefore is a concern for all practitioners.

How do we define fairness?

The tension among competing definitions and the impossibility of satisfying them all.

Individual Measures of Fairness

"Treat similar individuals similarly" [1]

Challenges:

- Definitions of "similar" are highly task dependent [2]
- Breaks down at decision boundaries [3]

[1] C. Dwork, M. Hardt, T. Pitassi, O. Reingold and R. Zemel, "Fairness Through Awareness," *Proceedings of the 3rd innovations in theoretical computer science conference*, pp. 214-226, 2012.
[2] A. Chouldechova and A. Roth, "The frontiers of fairness in machine learning," arXiv preprint arXiv:1810.08810, 2018.
[3] A. Narayanan, "21 Fairness Definitions and Their Politics," in Proc. Conf. Fairness, Accountability Transp., New York, 2018.

Group-based Fairness Measures

1. "Statistical Parity"
   - Example: admitting the same proportion of male applicants as female applicants to a college program.

2. "Error Rate Balance"
   - Example: ensuring the same rate of "false positive" results across racial or gender groups

3. "Predictive Parity"
   - Example: Accuracy of predictions are equal across groups

Challenge:

- It is impossible to achieve all three in the cases we usually care about! (Proven by Alexandra Chouldechova of CMU)

A. Chouldechova, "Fair prediction with disparate impact: A study of bias in recidivism prediction instruments.," Big Data, pp. 153-163, 2017.

Example: Courtroom algorithm accused of racial bias

Computer generated risk assessments were used to set bail, and for sentencing guidelines.

PROPUBLICA

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

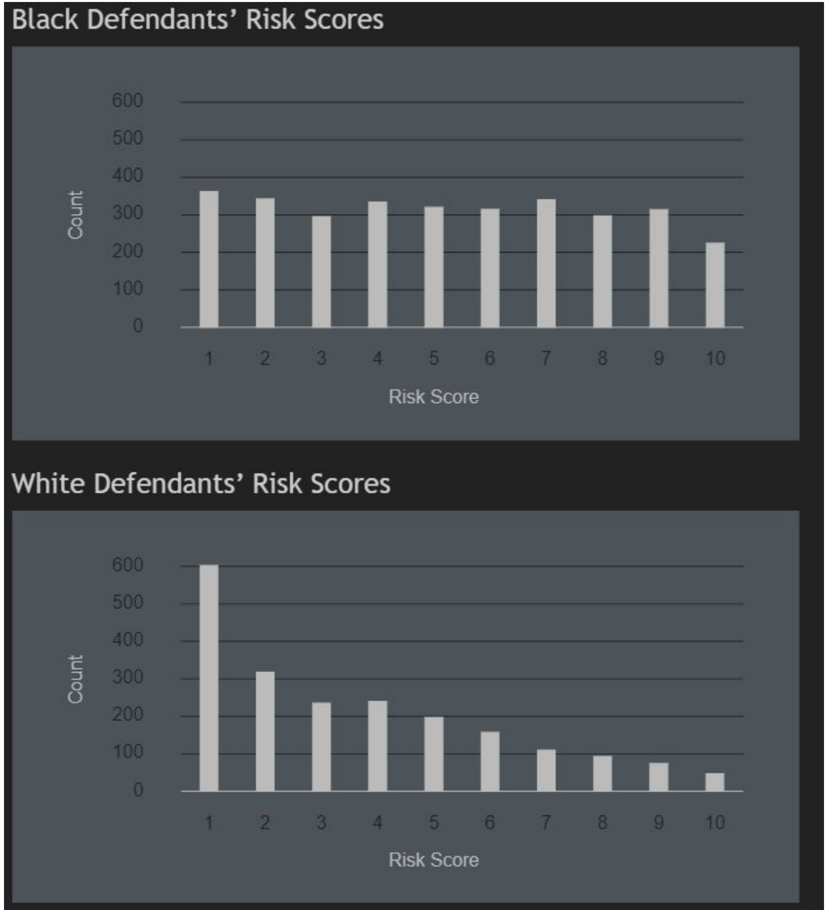
J. Angwin, J. Lason, S. Mattu and L. Kirchner, "Machine Bias," *ProPublica*, 23 May 2016.

Investigators: White defendants given lower risk scores

“Scores for white defendants were skewed toward lower-risk categories. Scores for black defendants were not.”

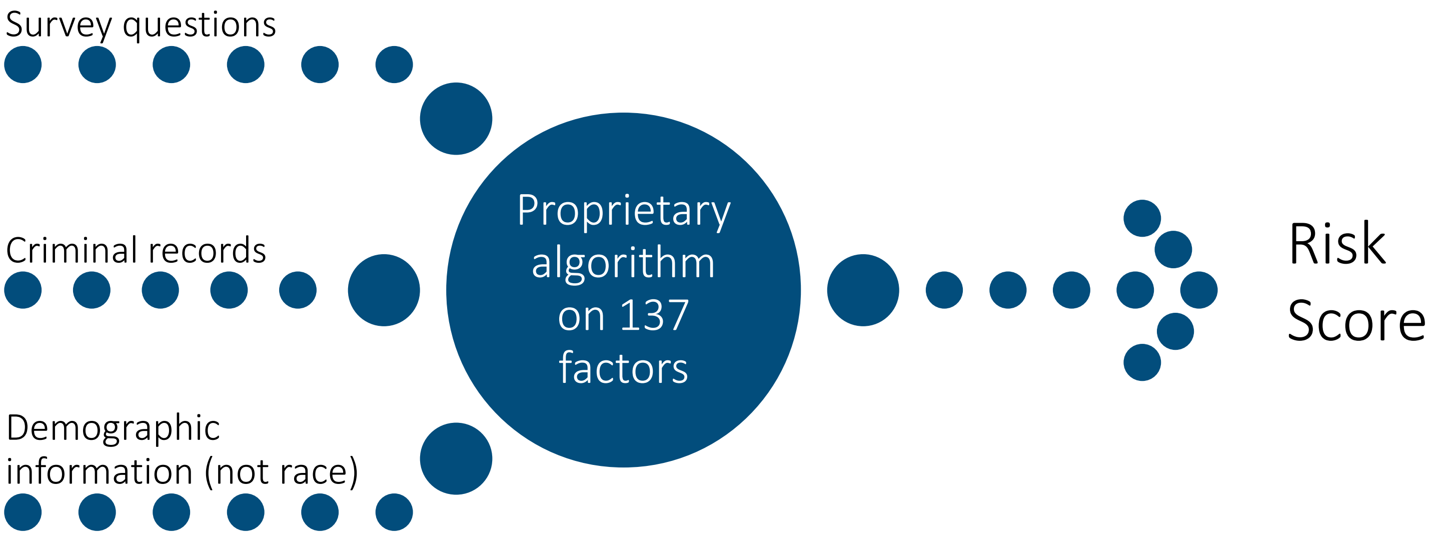
Source: ProPublica analysis of data from Broward County, Fla.

J. Angwin, J. Lason, S. Mattu and L. Kirchner, "Machine Bias," *ProPublica*, 23 May 2016.



These charts show that scores for white defendants were skewed toward lower-risk categories. Scores for black defendants were not. *(Source: ProPublica analysis of data from Broward County, Fla.)*

Investigators claim secret algorithm hides bias



Similarly, could a health risk score hide bias?

Is your model fair?

That can be a hard question to answer.

Whether an algorithm is unfair can be subjective

- ProPublica used “error rate balance” to accuse the COMPAS algorithm of racial bias.
- The makers of COMPAS used “predictive parity” to defend their algorithm as unbiased [1]

Who is right?

It depends on your definition of “fairness.”

[1] W. Dieterich, C. Mendoza and T. Brennan, "COMPAS risk scales: Demonstrating accuracy equity and predictive parity.," Northpointe, Inc., 2016.

Key idea:

2020 PREDICTIVE ANALYTICS 4.0 VIRTUAL EVENT

Machine learning fairness has no universal definition that applies to all situations. Attempts to satisfy all definitions of fairness are futile.

SOCIETY OF ACTUARIES®

So what do we do about machine learning fairness?

Actuaries are used to dealing with these issues

- Actuarial justification or maximal predictive accuracy vs. public policy goals or societal values
- Discussion around ACA age cost curve is a good illustration of this tension

Actuaries are used to dealing with these issues

Actuarial practice evolves to reflect societal values and public policy goals, like abolishing the use of race in mortality tables.

MORTALITY TABLE FOR U. S. WHITE MALES 1959-61

x	q_x	l_x	d_x	x	q_x	l_x	d_x
0	.02592	100,000	2,592	40	.00332	92,427	306
1	.00153	97,408	149	41	.00368	92,121	339
2	.00101	97,259	99	42	.00409	91,782	376
3	.00081	97,160	78	43	.00454	91,406	415
4	.00069	97,082	67	44	.00504	90,991	458
5	.00062	97,015	60	45	.00558	90,533	505

Image: C.W. Jordan, *Life Contingencies*. Chicago: Society of Actuaries, 1991.

Study offers example of combatting algorithmic bias

- Fundamental problem causing the bias was the choice of label.

RESEARCH ARTICLE

Science

ECONOMICS

Dissecting racial bias in an algorithm used to manage the health of populations

Ziad Obermeyer[1,2*], Brian Powers[3], Christine Vogeli[4], Sendhil Mullainathan[5*†]

- Model predicted healthcare expenditure, which turned out to be a biased proxy for underlying health status.
- Diagnosing the problem was possible because the algorithm manufacturer cooperated with researchers.

Z. Obermeyer, et al. "Dissecting racial bias in an algorithm used to manage the health of populations." *Science* 366.6464 (2019): 447-453.

Study offers example of combatting algorithmic bias

- A new label was developed: a basket of health outcomes and cost outcomes

- Measurable racial bias in model scoring was **reduced by 84%**

Z. Obermeyer. “Algorithms are as good as their labels,” Machine Learning for Healthcare Conference, Aug 3, 2020. Available: https://youtu.be/xt_pwq4HZWA.

How do machine learning practitioners avoid unintended bias?

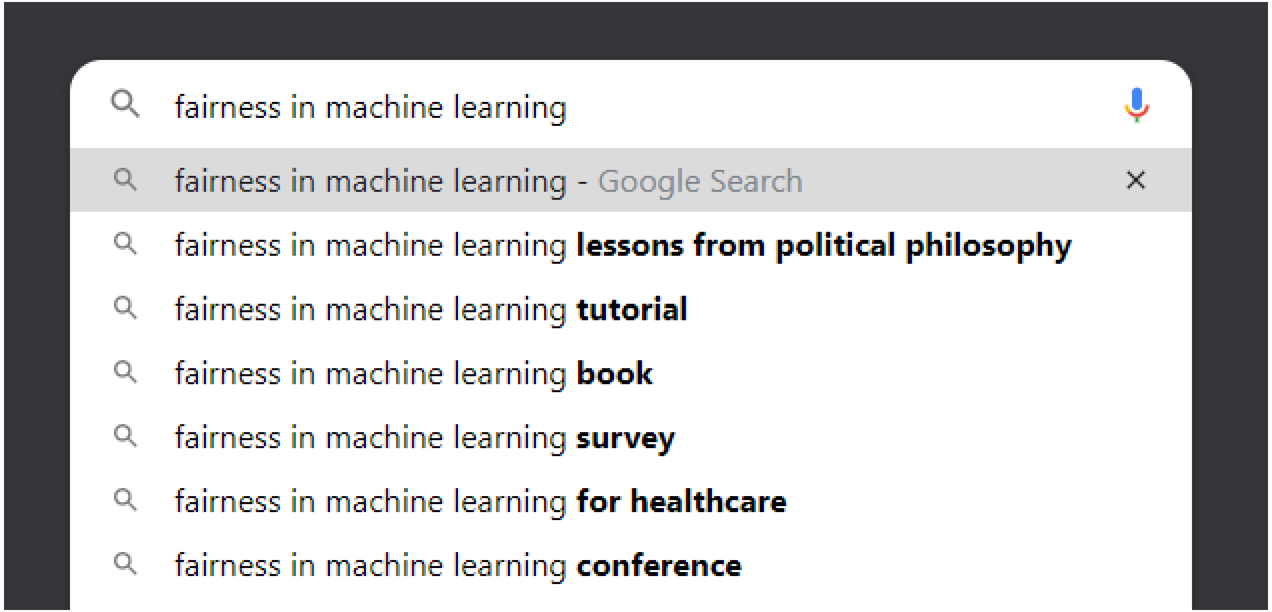
- First step: Awareness

Obliviousness to the problem is the first risk factor to overcome!

Seeking outside help to avoid algorithmic bias

- Fair ML conferences and publications
- Fairness audits offered as services from consulting companies
- Open source software packages with built-in fairness testing assistance

Without endorsing any specific product or service, a search will turn up a wealth of research and resources.



2020 PREDICTIVE ANALYTICS 4.0 VIRTUAL EVENT

Key Idea:

Machine learning fairness is a challenging issue, and it may be wise to consult outside resources for help.

SOCIETY OF ACTUARIES®

Concluding thoughts

The actuarial profession has a valuable role to play in the field of algorithmic fairness research and its related public policy debates.

2020 PREDICTIVE ANALYTICS 4.0 VIRTUAL EVENT

Questions?

Contact

Jonathan Culbert, Principal Data Scientist
Delta Dental of Michigan, Ohio, and Indiana
jculbert@deltadentalmi.com

SOCIETY OF ACTUARIES®

SOCIETY OF ACTUARIES®