
WHAT WILL BECOME OF US
↳ HOW TECHNOLOGY IS
CHANGING WHAT IT MEANS TO BE HUMAN.
NOVEMBER 18, 2018

"AI" is a Lie: Getting to the Real Issues

Momin M. Malik, PhD <momin_malik@cyber.harvard.edu>
Data Science Postdoctoral Fellow
Berkman Klein Center for Internet & Society at Harvard University

AGTech Forum, 13 December 2018

Slides: https://mominmalik.com/agtech2018.pdf

3-point summary

- Rule of thumb: Any place you see "AI" or "machine learning," substitute with "statistics"
- And any time you see "X predicts Y," read, "X correlates with Y"
- Only real-world testing (not simulated testing, nor real-world deployment) will tell if correlations will predict

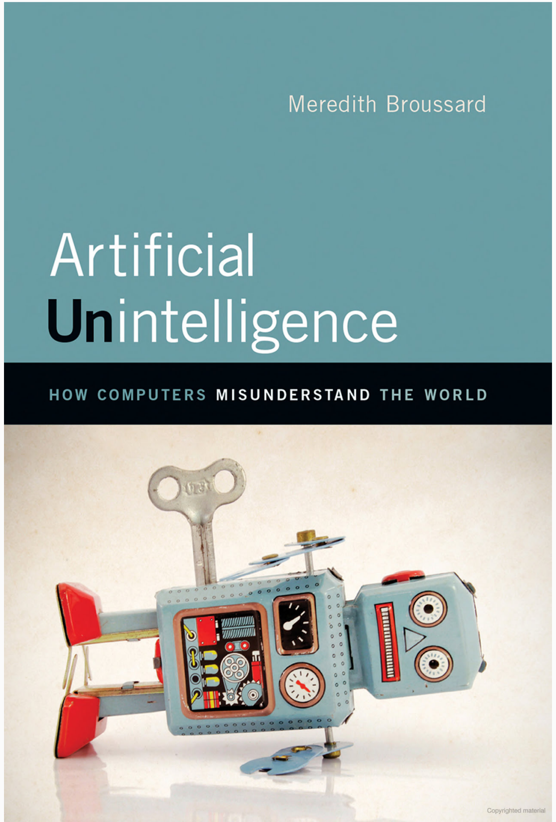
BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

About me

- DEPARTMENT OF THE HISTORY OF SCIENCE, HARVARD UNIVERSITY
- Berkman — The Berkman Center for Internet & Society at Harvard University
- OXFORD INTERNET INSTITUTE — UNIVERSITY OF OXFORD
- Carnegie Mellon University School of Computer Science
- BERKMAN KLEIN CENTER FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

“AI” is not AI, but statistics

"So, it's not real AI?"



- "So, it's not real AI?" he asked.
- "Oh, it's real," I said. "And it's spectacular. But you know, don't you, that there's no simulated person inside the machine? Nothing like that exists. It's computationally impossible."
- His face fell. "I thought that's what AI meant," he said. "I heard about IBM Watson, and the computer that beat the champion at Go, and self-driving cars. I thought they invented real AI."

Internally, joking about mismatch

Baron Schwartz @xaprb

Follow

When you're fundraising, it's AI
When you're hiring, it's ML
When you're implementing, it's linear regression
When you're debugging, it's printf()

12:52 AM - 15 Nov 2017

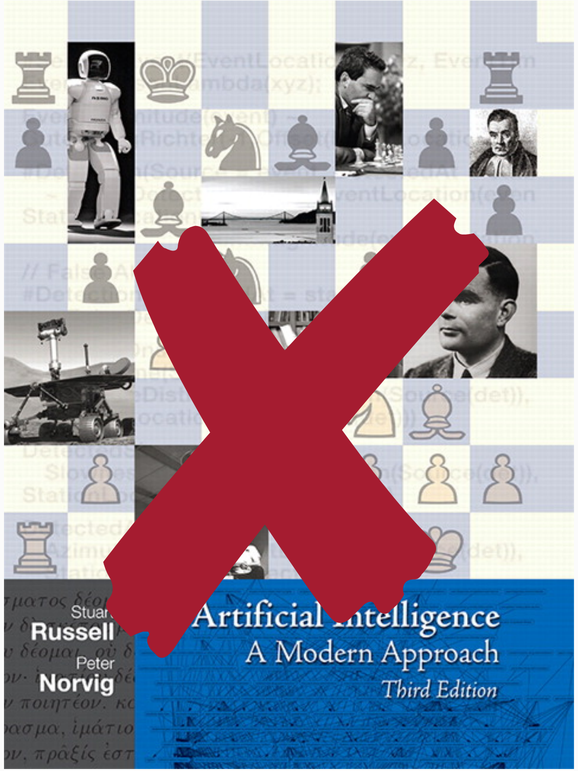
5,545 Retweets **12,654** Likes

90 5.5K 13K

xkcd.com/1838

AI changed, but kept the same name!

- "After about 14 years of trying to get language models to work using logical rules, I started to adopt probabilistic approaches." (Peter Norvig, "On Chomsky," 2010)



Aspirational naming misleads

ARTIFICIAL INTELLIGENCE MEETS NATURAL STUPIDITY

Drew McDermott
MIT AI Lab Cambridge, Mass 02139

As a field, artificial intelligence has always been on the border of respectability, and therefore on the border of crackpottery. Many critics <Dreyfus, 1972>, <Lighthill, 1973> have urged that we are over the border. We have been very defensive toward this charge, drawing ourselves up with dignity when it is made and folding the cloak of Science about us. On the other hand, in private, we have been justifiably proud of our ... ideas, because pursuing them is the onl...

Unfortunately, the necessity for s... the culture of the hacker in computer ... to cripple our self-discipline. In a young field, self-discipline is not necessarily a virtue, but we are not getting any younger. In the past few years, our tolerance of sloppy thinking has led us to repeat many mistakes over and over. If we are to retain any credibility, this should stop.

This paper is an effort to ridicule some of these mistakes. Almost everyone I know should find himself the target at some point or other; if you don't, you are encouraged to write up your own favorite fault. The three described here I suffer from myself. I hope self-ridicule will be a complete catharsis, but I doubt it. Bad ... though, if we can't ... ple. ... m crowd for

Wishful Mnemonics

Wishful Mnemonics

A major source of simple-mindedness in AI programs is the use of mnemonics like "UNDERSTAND" or "GOAL" to refer to programs and data structures. This practice has been inherited from more

Compare the mnemonics in Planner <Hewitt,1972> with those in Conniver <Sussman and McDermott, 1972>:

Planner	Conniver
GOAL	FETCH & TRY-NEXT
CONSEQUENT	IF-NEEDED
ANTECEDENT	IF-ADDED
THEOREM	METHOD
ASSERT	ADD

It is so much harder to write programs using the terms on the right! When you say (GOAL . . .), you can just feel the enormous power at your fingertips. It is, of course, an illusion.

When you say (GOAL . . .), you can just feel the enormous power at your fingertips. It is, of course, an illusion.

... 1965> What if atomic symbols had been called "concepts", or CONS had been called ASSOCIATE? As it is, the programmer has no debts to pay to the system. He can build whatever he likes. There are some minor faults; "property lists" are a little risky; but by now the term is sanitized.

Resolution theorists have been pretty good about wishful mnemonics. They thrive on hitherto meaningless words like RESOLVE and PARAMODULATE, which can only have their humble, technical meaning. There are actually quite few pretensions in the resolution literature. <Robinson, 1965> Unfortunately, at the top of their intellectual edifice stand the word "deduction". This is *very* wishful, but not entirely their fault. The logicians who first misused the term (e.g., in the "deduction" theorem) didn't have our problems; pure resolution theorists don't either. Unfortunately, too many AI researchers took them at their word and assumed that deduction, like payroll processing, had been tamed.

Of course, as in many such cases, the only consequence in the long run was that "deduction" changed in meaning, to become something narrow, technical, and not a little sordid.

Page 4 SIGART Newsletter No. 57 April 1976

"Prediction" is not prediction, but retrospective

Prediction seems scary powerful

MIT Technology Review

Topics+ The Download Magazine Events

Intelligent Machines

Software Predicts Tomorrow's News by Analyzing Today's and Yesterday's

Prototype software can give early warnings of disease or violence outbreaks by spotting clues in news reports.

by Tom Simonite February 1, 2013

A method of using online information to accurately predict the future could transform many industries.

BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY
AT HARVARD UNIVERSITY

Predict... the future?

Predicting the Future With Social Media

Sitaram Asur
Social Computing Lab
HP Labs
Palo Alto, California
Email: sitaram.asur@hp.com

Bernardo A. Huberman
Social Computing Lab
HP Labs
Palo Alto, California
Email: bernardo.huberman@hp.com

Mar 2010

***Abstract*—In recent years, social media has become ubiquitous and important for social networking and content sharing. And yet, the content that is generated from these websites remains largely untapped. In this paper, we demonstrate how social media content can be used to predict real-world outcomes. In particular, we use the chatter from Twitter.com to forecast box-office revenues for movies. We show that a simple model built from**

This paper reports on such a study. Specifically we consider the task of predicting box-office revenues for movies using the chatter from Twitter, one of the fastest growing social networks in the Internet. Twitter [1], a micro-blogging network, has experienced a burst of popularity in recent months leading to a huge user-base, consisting of several tens of millions of

Predicting the Future — Big Data, Machine Learning, and Clinical Medicine

Ziad Obermeyer, M.D., and Ezekiel J. Emanuel, M.D., Ph.D.

By now, it's almost old news: big data will transform medicine. It's essential to remember, however, that data by themselves are useless. To be useful, data must be analyzed, interpreted, and acted on. Thus, it is algorithms — not data sets — that will prove transformative. We believe, therefore, that attention has to shift to new statistical tools from the field of machine learning that will be critical for anyone practicing medicine in the 21st century.

First, it's important to understand what machine learning is not. Most computer-based algorithms in medicine are "expert systems" — rule sets encoding knowledge on a given topic, which are applied to draw conclusions

1216 N ENGL J MED 375;13 NEJM.ORG SEPTEMBER 29, 2016

Merriam-Webster SINCE 1828

predict *verb*

pre·dict | \pri-ˈdikt \

predicted; predicting; predicts

Definition of *predict*

transitive verb

: to declare or indicate in advance

especially : foretell on the basis of observation, experience, or scientific reason

intransitive verb

: to make a prediction

Other Words from *predict*

Synonyms

Choose the Right Synonym

Prediction is not prediction

- "*It's not prediction at all!* I have not found a single paper predicting a future result. All of them claim that a prediction could have been made; i.e. they are *post-hoc* analysis and, needless to say, negative results are rare to find." (Daniel Gayo-Avello, 2012)

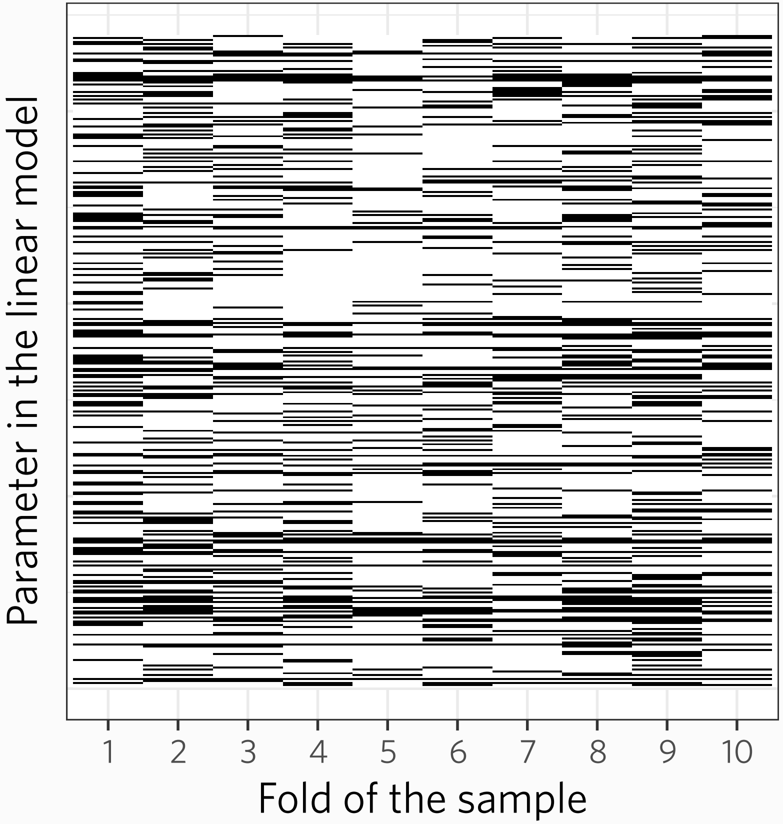
Prediction is correlation, not causation

BERKMAN KLEIN CENTER FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

Prediction is correlation

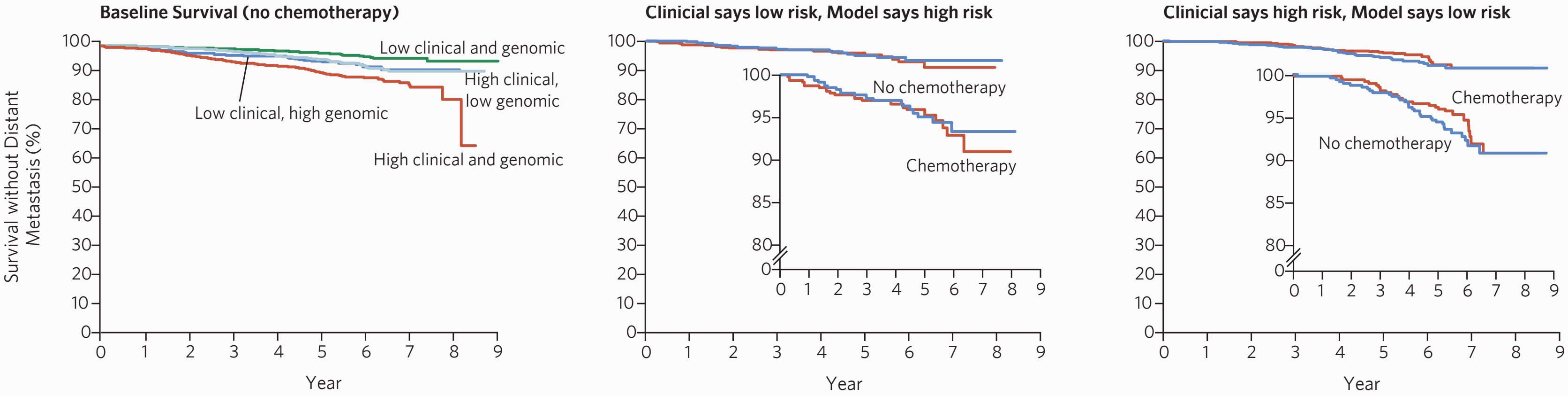
- Spurious (non-causal) correlations can fit the data really well!
- Google Flu Trends: half flu detector, half winter detector

Intervention requires causality



- Very different sets of correlations can "predict" equally well
- But they suggest very different interventions

At a minimum, demand real-world experimental testing

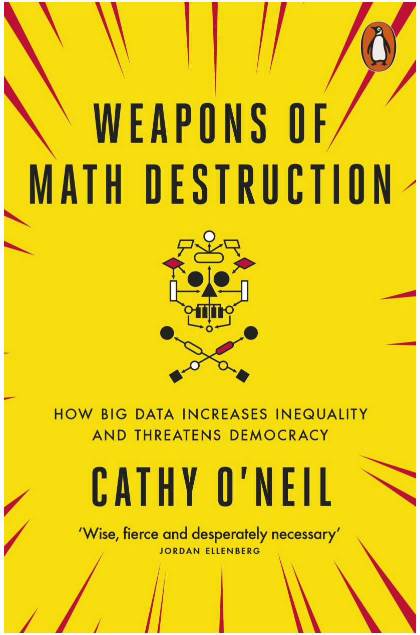


- Breast cancer: when a machine learning model said "high risk" but clinical risk was low, chemo made things worse!
- (But can help avoid unnecessary chemo)

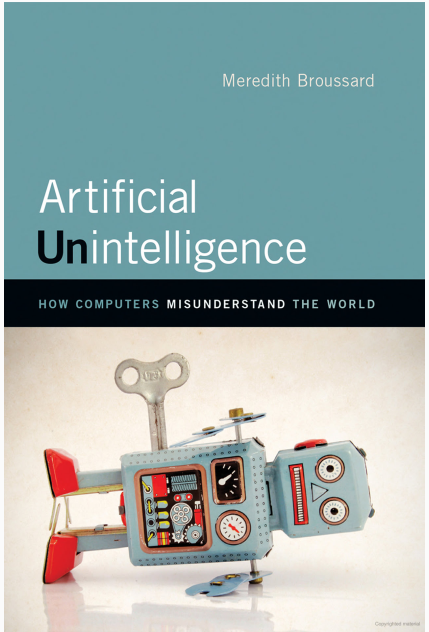
Summary

- There's lots of misleading language. Don't believe the hype, or everything you hear.
- Anything using "artificial intelligence" or "machine learning" is going to be statistical
- AI, ML are based on *correlations*. Among other issues, they can go wrong in every way that correlations can go wrong.

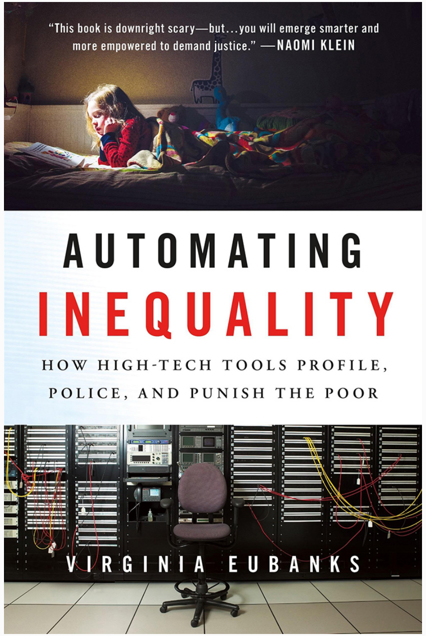
Further reading



- One of the earliest, still one of the best!



- **Chapter 7: best machine learning overview for lay audience**
- (But, two subtle mistakes: see mominmalik.com/broussard)



- Stories showing that *implementation is key*. The best intentions, and most careful technical work, can go awry.


WHAT WILL BECOME OF US
↳ HOW TECHNOLOGY IS
CHANGING WHAT IT MEANS TO BE HUMAN.
NOVEMBER 18, 2018

BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

References/Credits

Introduction

Robot holding skull: Cover image of "What Will Become of Us?", *New York Times Magazine* (The Tech & Design issue), 14 November 2018. Concept by delcan & company. Photo illustration by Jamie Chung. Prop styling by Pink Sparrow. C.G. work by Justin Metz. https://www.nytimes.com/2018/11/14/magazine/behind-the-cover-what-will-become-of-us.html.

Terminator skull: Nemesis Now Ltd, Terminator Skull Box T-800 (18CM). https://www.menkind.co.uk/terminator-t800-skull-box.

Hand: Hayati Kayhan, Holding human skull in hand, Conceptual image (Shakespeare's Hamlet scene concept). 19 October 2014.

"AI" is not AI

Stuart Russel and Peter Norvig, *Artificial Intelligence: A Modern Approach* (3rd Edition). Pearson, 2009.

Peter Norvig, "On Chomsky and the Two Cultures of Statistical Learning." http://norvig.com/chomsky.html.

Larry Wasserman, "Rise of the Machines." In *Past, Present and Future of Statistical Science*, ed. by Xihong Lin, Christian Genest, David Banks, Geert Molenberghs, David Scott, and Jane-Ling Wang, 525–536. CRC Press, 2014.

Drew McDermott, "Artificial Intelligence meets Natural Stupidity." *SIGART Bulletin* 57 (April 1976): 4–9.

"Prediction" is not prediction

Tom Simonite, "Intelligent Machines: Software Predicts Tomorrow's News by Analyzing Today's and Yesterday's." *MIT Technology Review*, 1 February 2013. https://www.technologyreview.com/s/510191/software-predicts-tomorrows-news-by-analyzing-todays-and-yesterdays.

Sitaram Asur and Bernando A. Huberman, "Predicting the Future With Social Media." In *Proceedings of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology* (WI-IAT '10), 492–499. 2010. https://dx.doi.org10.1109/WI-IAT.2010.63.

Ziad Obermeyer and Ezekiel J. Emanuel, "Predicting the Future — Big Data, Machine Learning, and Clinical Medicine." *New England Journal of Medicine* 375, no. 13 (2016): 1216–1219. https://dx.doi.org/10.1056/NEJMp1606181.

Merriam-Webster. "predict". Merriam-Webster.com, 2018. https://www.merriam-webster.com/dictionary/predict.

Daniel Gayo-Avello, "'I Wanted to Predict Elections with Twitter and all I got was this Lousy Paper': A Balanced Survey on Election Prediction using Twitter Data." *arXiv*, 28 April 2012. https://arxiv.org/abs/1204.6441.

Daniel Gayo-Avello, "No, You Cannot Predict Elections with Twitter." *IEEE Internet Computing* 16 (2012): 91–94. https://dx.doi.org/10.1109/MIC.2012.137.

"Prediction" is not causation

David Lazer, Ryan Kennedy, Gary King, and Alessandro Vespignani, "The Parable of Google Flu Trends: Traps in Big Data Analysis." *Science* 343 (14 March 2014): 1203–1205. https://dx.doi.org/10.1126/science.1248506.

Sendhil Mullainathan and Jan Spiess, "Machine Learning: An Applied Econometric Approach." *Journal of Economic Perspectives* 31, no. 2 (2017): 87–106. https://dx.doi.org/10.1257/jep.31.2.87.

Fatima Cardoso, Laura J. van't Veer, Jan Bogaerts, Leen Slaets, Giuseppe Viale, Suzette Delaloge, Jean-Yves Pierga, Etienne Brain, Sylvain Causeret, Mauro DeLorenzi, Annuska M. Glas, Vassilis Golfinopoulos, et al., "70-Gene Signature as an Aid to Treatment Decisions in Early-Stage Breast Cancer." New England Journal of Medicine 375 (25 August 2016): 717-729. https://dx.doi.org/10.1056/NEJMoa1602253.

Summary

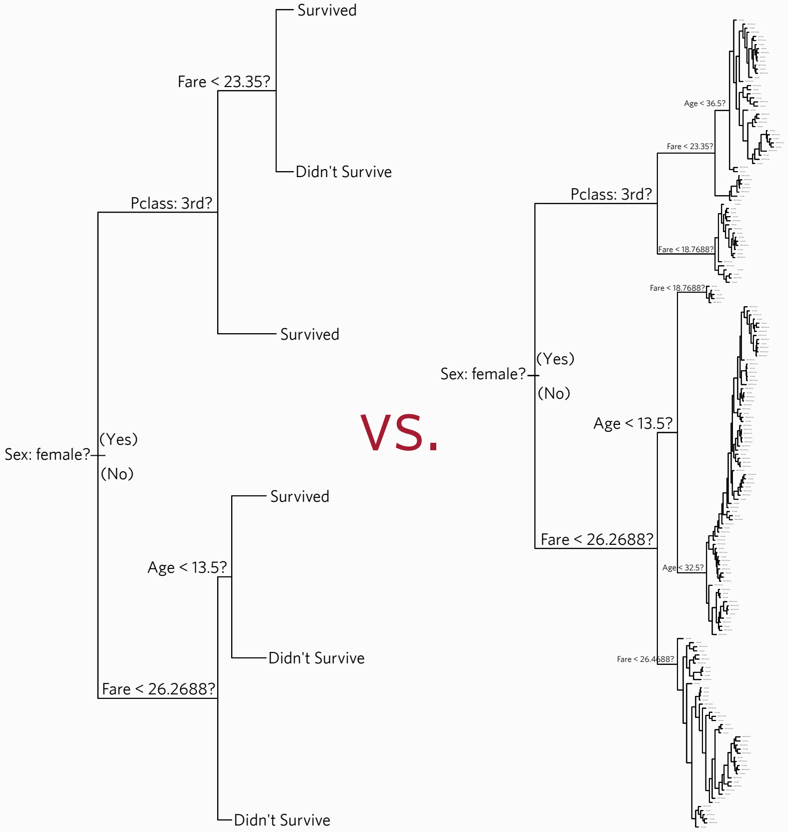
Cathy O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown, 2016.

Meredith Broussard, *Artificial Unintelligence: How Computers Misunderstand the World*. MIT Press, 2018.

Virginia Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. St. Martin's Press, 2018

Correlations can “overfit”

- *Overfitting*: Model finds correlation to random ‘noise’ (“memorizes the data”)
- (Overfitting is simpler than “*p*-hacking,” but similar)
- Existing solution: split the data (e.g., 1:1, 4:1, 9:1). “Hold out” one part. Idea: *The signal should be the same, but not the noise.* Testing on held out data (cross-validation) should reveal overfitting



But cross-validation can fail

- Re-using the test set can overfit to the test set! E.g., Kaggle competitions
- Or, if there are dependencies (temporal, network, group) between data splits, it "shares" information
- E.g., temporal: Fitting on values that come after test values is "time traveling"!