BERKMAN KLEIN CENTER FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

Correlates of oppression: Machine learning and society

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

MIT CMS.701/CMS.901: Current Debates in Media, 30 October 2019
Instructor: Dr. Sasha Costanza-Chock
Slides: https://mominmalik.com/cms2019.pdf

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

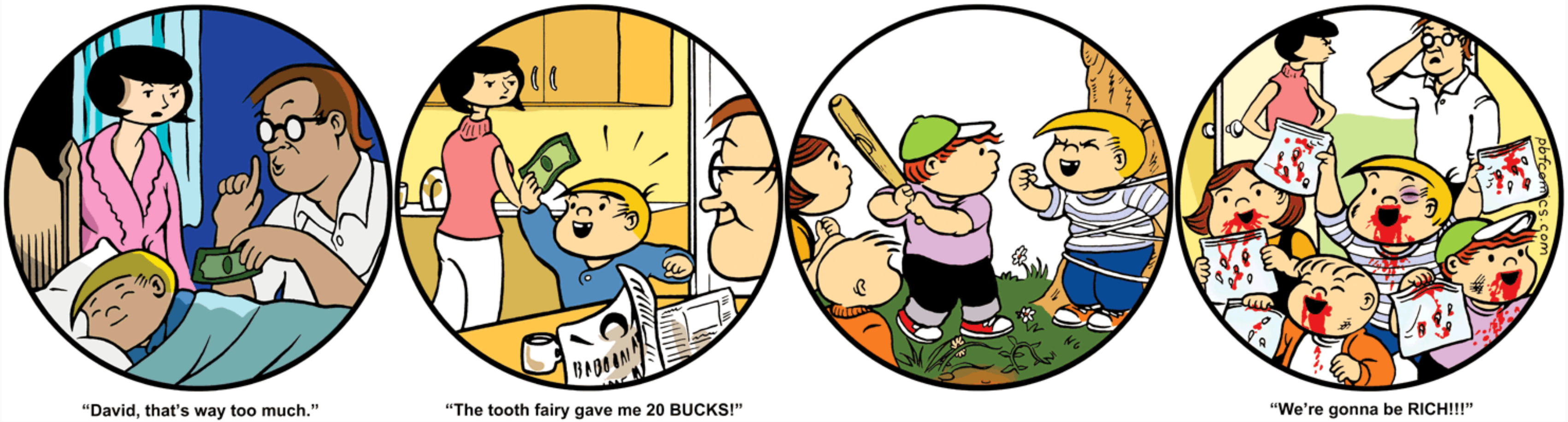
Learning objectives

- Recognize what substance there is (and isn't) in machine learning claims of "prediction"
- Understand the implications of using the *central tendency*
- Link oppressive possibilities to uses of machine learning as *top-down, imposed optimization* based on *surveillable proxies* for *intimate quantities*

Outline

- Prediction
- Statistics vs. machine learning
- Correlation and proxies
- Central tendency
- A positive example

"The essence of machine learning"



Prediction

Predict future behavior

MIT Technology Review

Topics+ The Download Magazine Events

The Download

What's up in emerging technology

April 13, 2018

Facebook is using AI to predict users' future behavior and selling that data to advertisers

In confidential documents seen by the *Intercept*, Facebook touts its ability to "improve" marketing outcomes with what it calls "loyalty prediction."

Newspeak: The AI software that powers this capability, called "FBLearner Flow," was first announced in... **Read more**

Predict birdsong

MIT Technology Review

Topics+ The Download Magazine Events

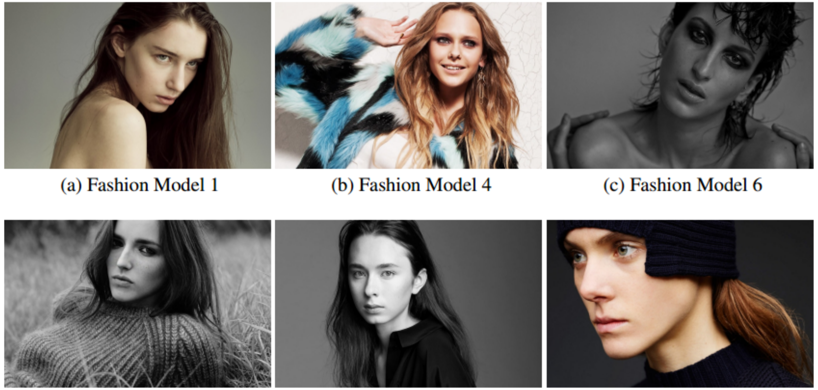
Rewriting Life

Scientists Can Read a Bird’s Brain and Predict Its Next Song

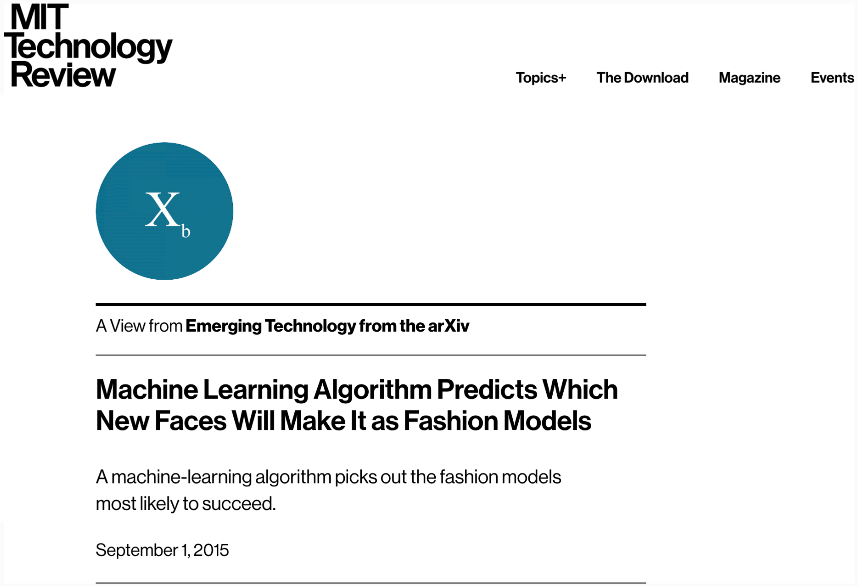
Next up, predicting human speech with a brain-computer interface.

by Antonio Regalado October 11, 2017

Predict fashion models success



(a) Fashion Model 1 (b) Fashion Model 4 (c) Fashion Model 6
(d) Fashion Model 7 (e) Fashion Model 8 (f) Fashion Model 9



MIT Technology Review

Topics+ The Download Magazine Events

A View from **Emerging Technology from the arXiv**

Machine Learning Algorithm Predicts Which New Faces Will Make It as Fashion Models

A machine-learning algorithm picks out the fashion models most likely to succeed.

September 1, 2015

Predict *news*

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, "I Wanted to Prediction Elections with Twitter and all I got was this Lousy Paper", 2012

Prediction is not prediction

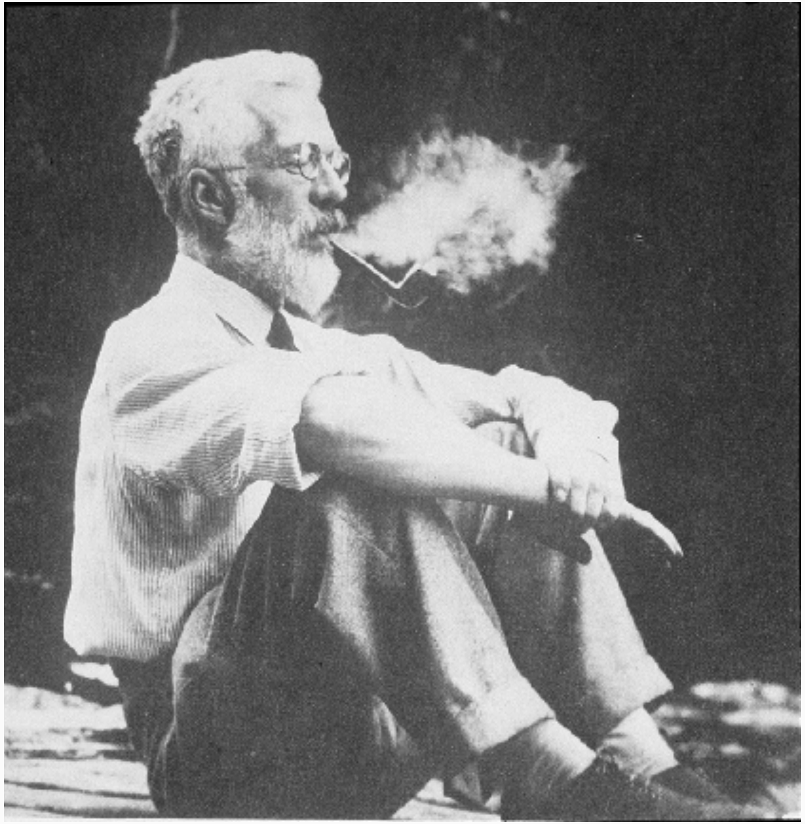
- Predictions are post-hoc $\boxed{\textit{correlations}}$
- Maybe these work as a basis for prediction, but maybe not
- Is *static* prediction: not prediction under *change*, nor prediction under *intervention*. Not an inevitable, or even natural, usage of "prediction" (Rescher, 1998)

Statistics vs. machine learning

BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

What is statistics?

- "briefly, and in its most concrete form, the object of statistical methods is the reduction of data."
  - R. A. Fisher, "On the Mathematical Foundations of Theoretical Statistics" (1922)
- This definition has nothing to do with probability!



What is machine learning?

- "A computer program is said to learn from experience *E* with respect to some class of tasks *T* and performance measure *P* if its performance at tasks in *T*, as measured by *P*, improves with experience *E*."
  – Tom M. Mitchell, *Machine Learning*, 1997
- This definition has nothing to do with probability, statistics, or even data!

14 years of failure with logical rules

- "As Steve Abney wrote in 1996, 'In the space of the last ten years, statistical methods have gone from being virtually unknown in computational linguistics to being a fundamental given.'... **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

Repackaged statistics

- "1980s–1990s work in machine learning often replayed insights available in traditional statistics... Indeed, it became increasingly clear through the 1990s that **many 'insights' of connectionism were differently named versions of statistical techniques**."

  – Maggie Boden, *Mind as Machine*, 2006

Transition "in a matter of a few years"

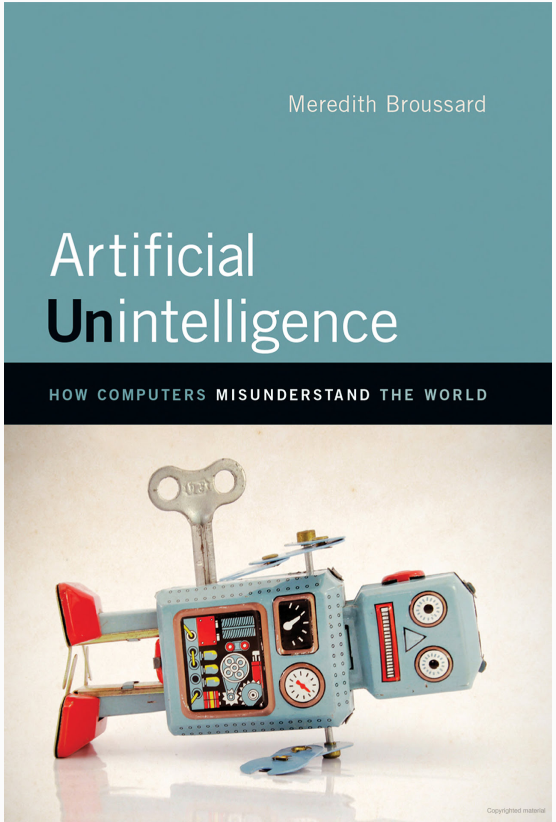
- "At first, ML researchers developed... a collection of rather primitive (yet clever) set of methods to do classification... that eschewed probability. But very quickly they adopted advanced statistical concepts like empirical process theory and concentration of measure. **This transition happened in a matter of a few years**."

  – Larry Wasserman, "Rise of the Machines", 2014

A "second culture" of statistics

- "In the past fifteen years, the growth in algorithmic modeling applications and methodology has been rapid. It has occurred largely outside statistics in **a new community—often called machine learning—that is mostly young computer scientists**."
  - Leo Breiman, "Statistical Modeling: The Two Cultures", 2001

Names are strategically misleading



- "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."

Names are strategically misleading

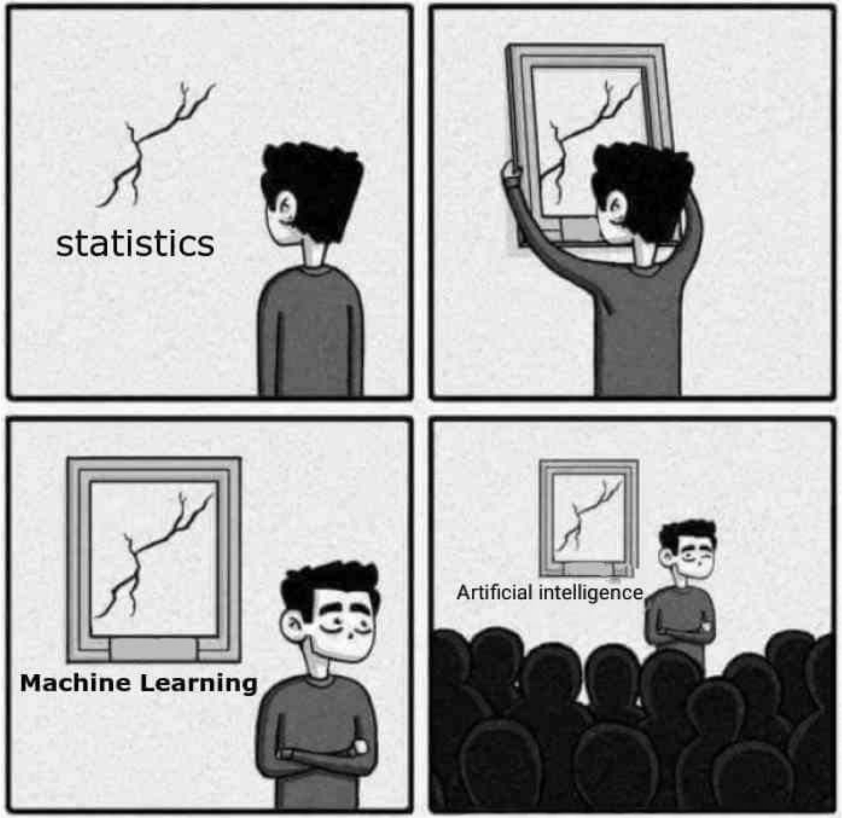
Baron Schwartz @xaprb
Follow

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

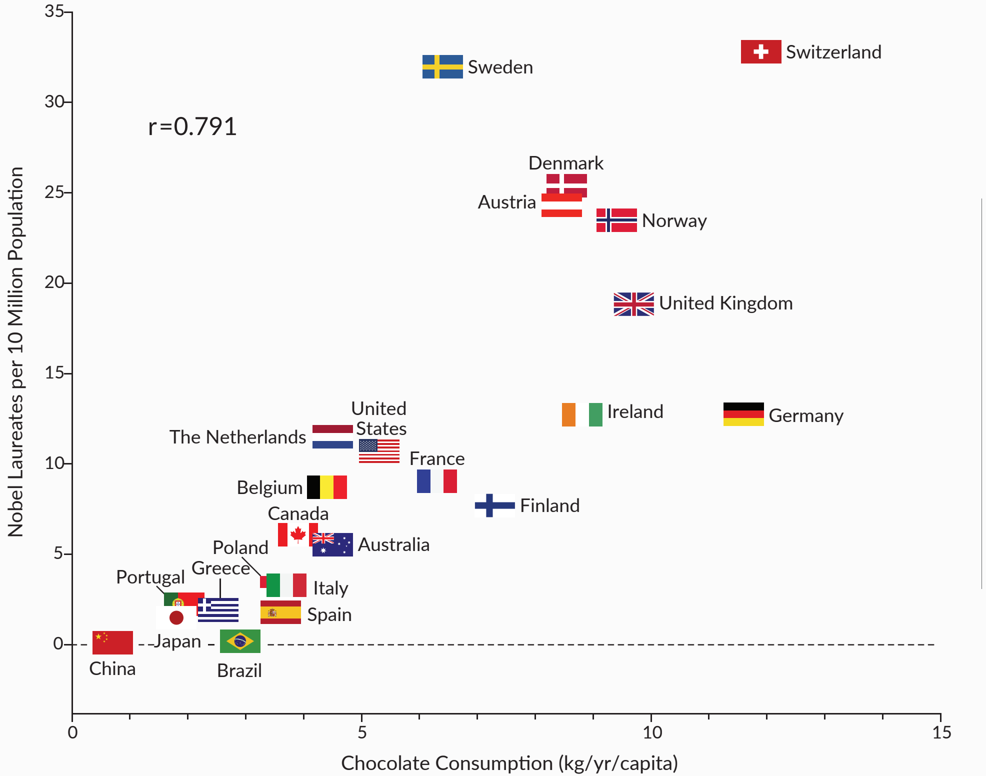


Still, are differences

- There is no "learning" (if there ever was) other than as metaphor
- ML started off doing "algorithms" and still uses that term, but the logic is really one of *statistical models*
- There are differences: statistics are used in ways anathema to traditional statisticians. No social theory, only optimal correlations. "Instrumentalist", "data positivism" (Jones, 2018)
- With this, ML found ways to get correlations in new types of data: images, audio, words

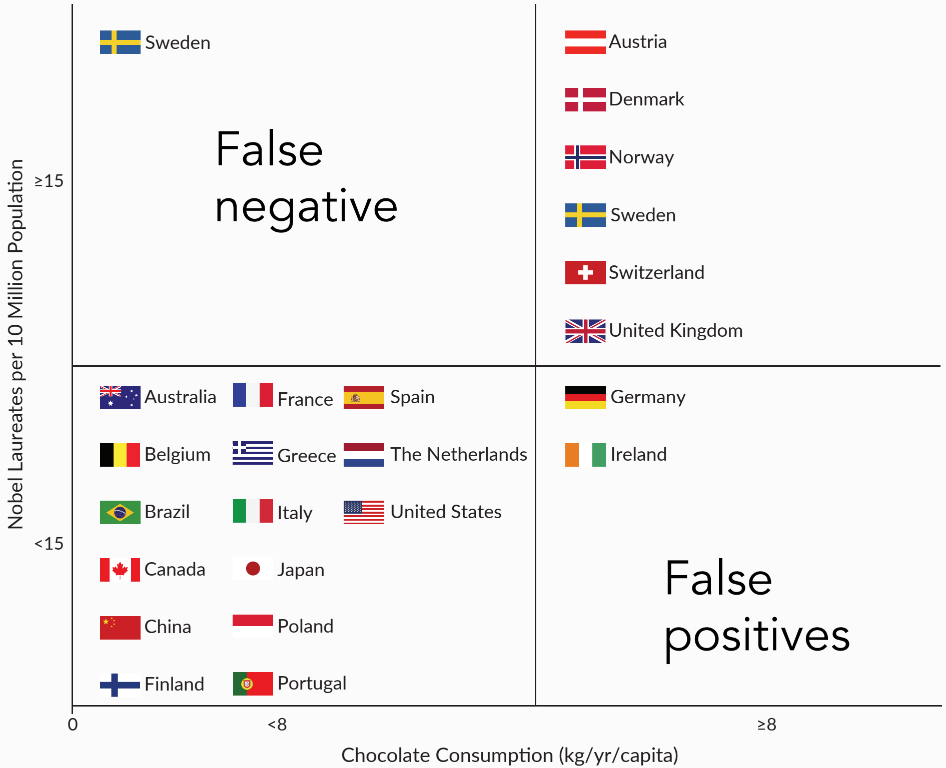
Correlation and proxies

Co-related: Nobel prizes and chocolate



(Messerli, 2012)

Discrete version: the "majority class"



BERKMAN
KLEIN CENTER
FOR INTERNET & SOCIETY
AT HARVARD UNIVERSITY

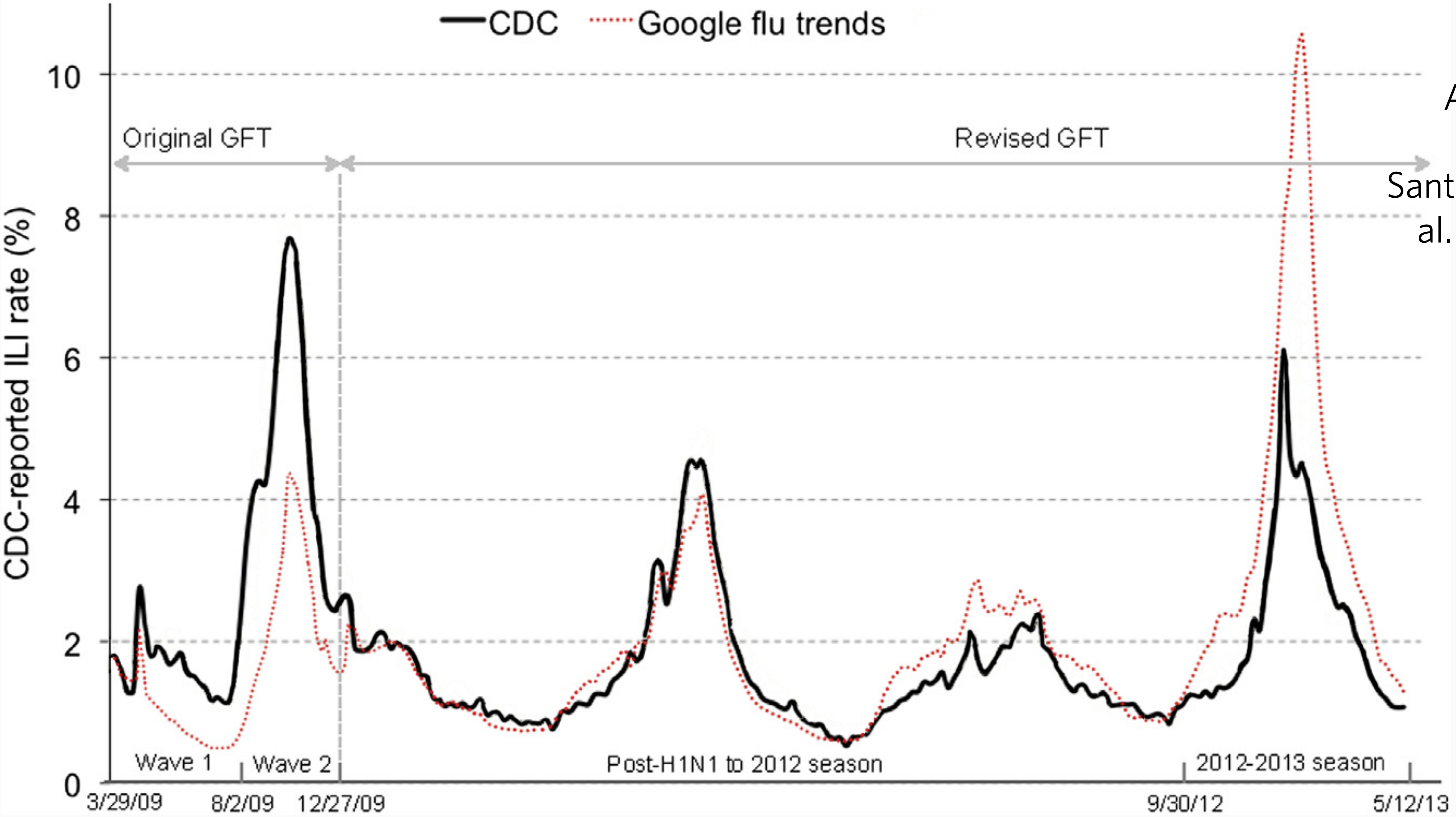
Are correlations enough to predict?

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

(Lazer et al., 2014)

Correlations can fail

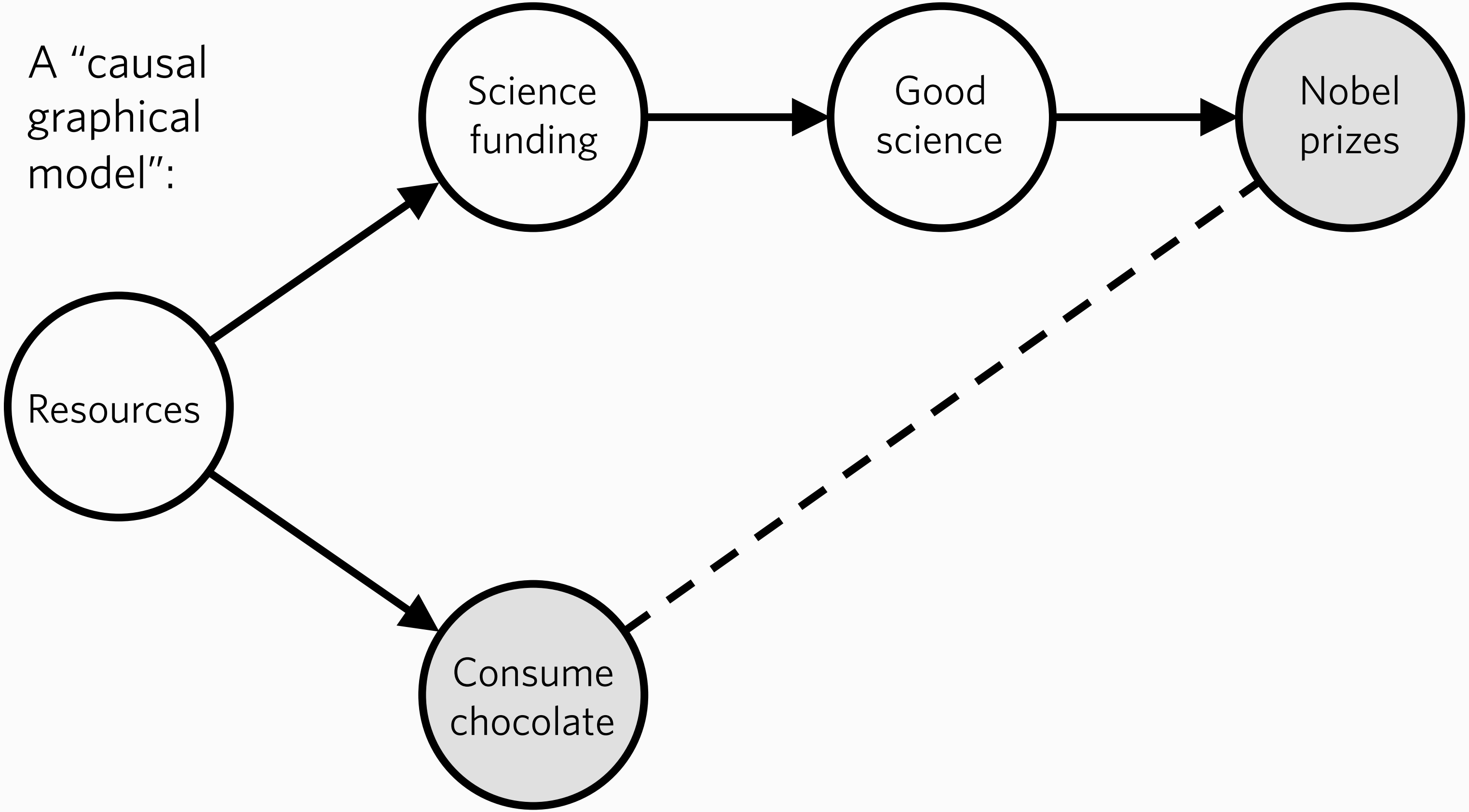
- Introduction
- Prediction
- Statistics vs. machine learning
- Correlation and proxies
- Central tendency
- A positive example
- Conclusion
- References



Adapted from Santillana et al. (2014).

Nobel prizes: *cause* is resources

A "causal graphical model":



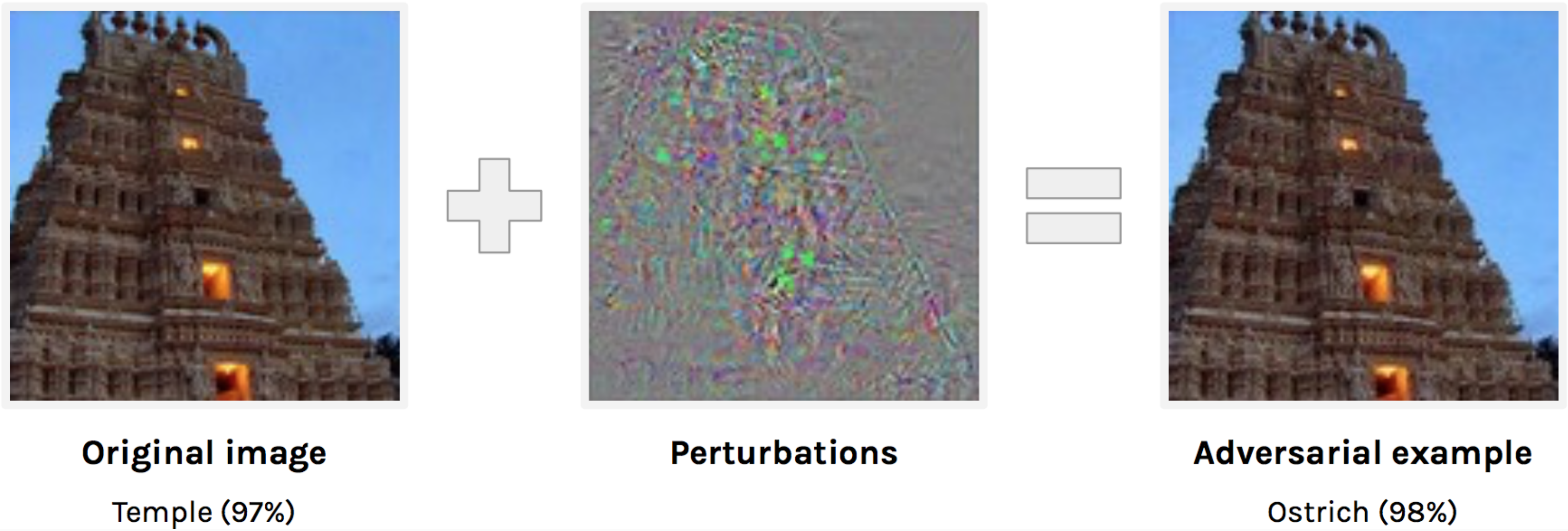
Correlations are proxies

- Much of what machine learning has accomplished is finding ways to find correlations
- E.g., between human labels and groups of pixels
- But proxies can always be gamed, which makes them fail ("McNamara's fallacy")
- And we need to know the target signal: in many cases, this must be laboriously, manually collected. "Automation's last mile" (Mary Gray and Siddharth Suri, *Ghost Work*, 2019)

Are proxies just?

- Is it okay to use optimal proxies?
- In the 1890s, life insurance companies charged more for African Americans, arguing that they had shorter life expectancy (Bouk, 2015)
- Car insurance is still a “ghetto tax” (Fergus, 2013)
- Example from Deborah Hellman: people who experience intimate partner violence have higher health insurance costs. Should we therefore charge them more?
- Feminist and civil rights campaigners in the 1970s argued for collectivizing risk; they lost to the insurance industry and “actuarial fairness” (Horan, 2011)

Gaming proxies



(Despois, 2017)

Gaming proxies in realtime

(Eykholt et al., 2018)

The *labor* of the target signal

"Source subject": Marquese Scott

Caroline Chan, "Everybody Dance Now: Motion Retargeting Video Subjects." https://youtu.be/PCBTZh41Ris

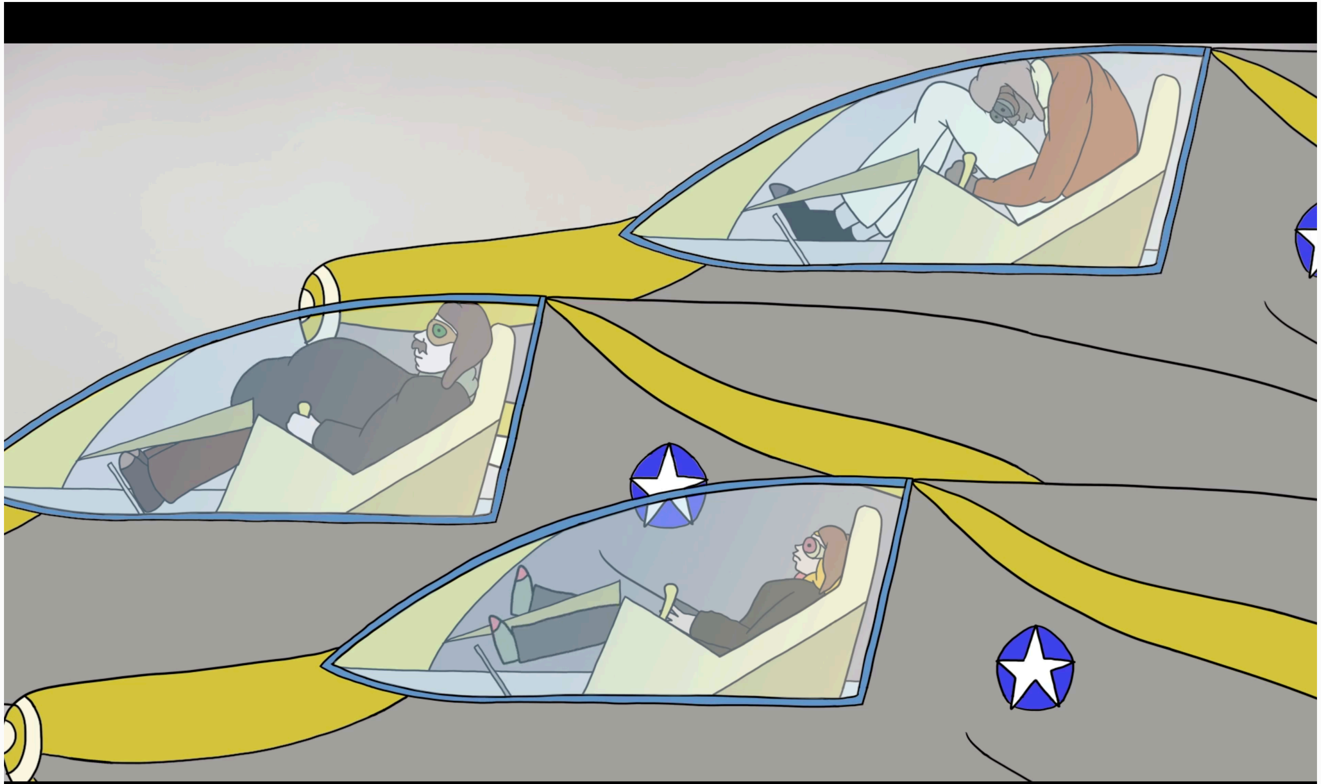
Central tendency

Correlations are a "central tendency"

The problem with central tendency

- Machine learning predictions are of a *central tendency*
- Historically, the idea of using central tendencies was seen as strange and highly contested!
- By choosing central tendency, we choose to punish outliers!
- Central tendency can never treat people as individuals. No real "personalized medicine", "individual risk scores"

The “flaw of averages”

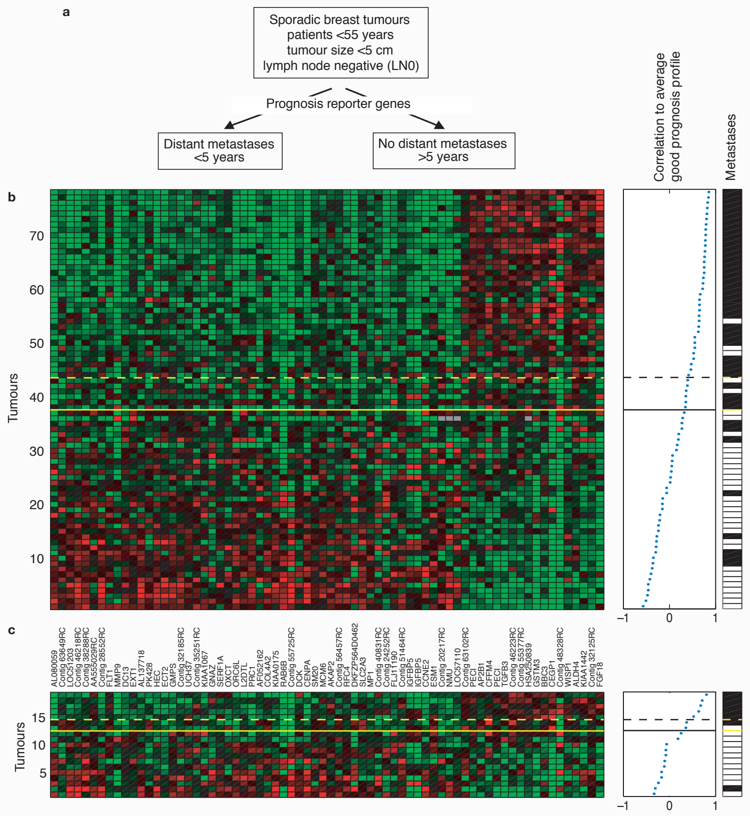


(Rose, 2014)

BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY
AT HARVARD UNIVERSITY

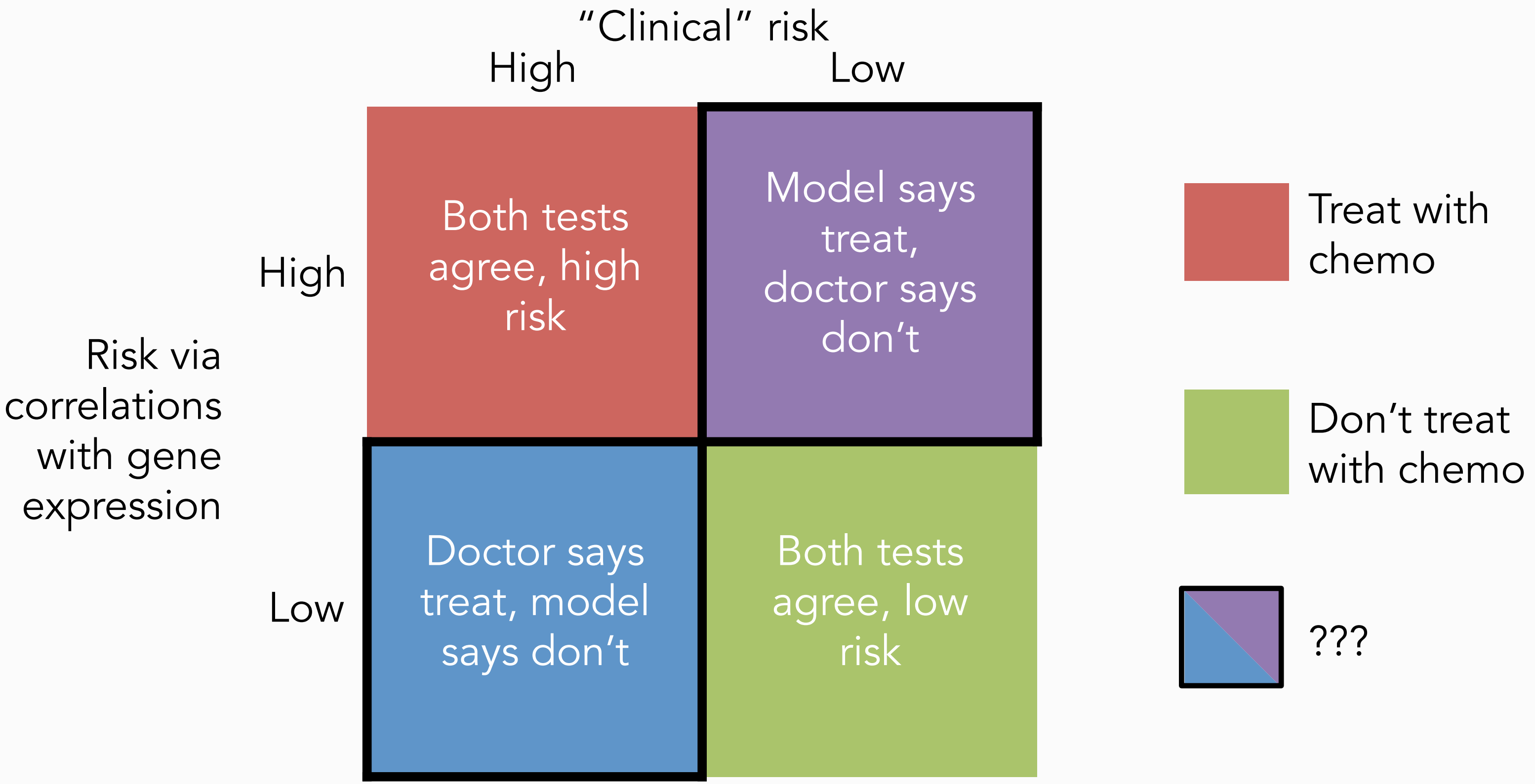
A positive example

Genes correlated with breast cancer



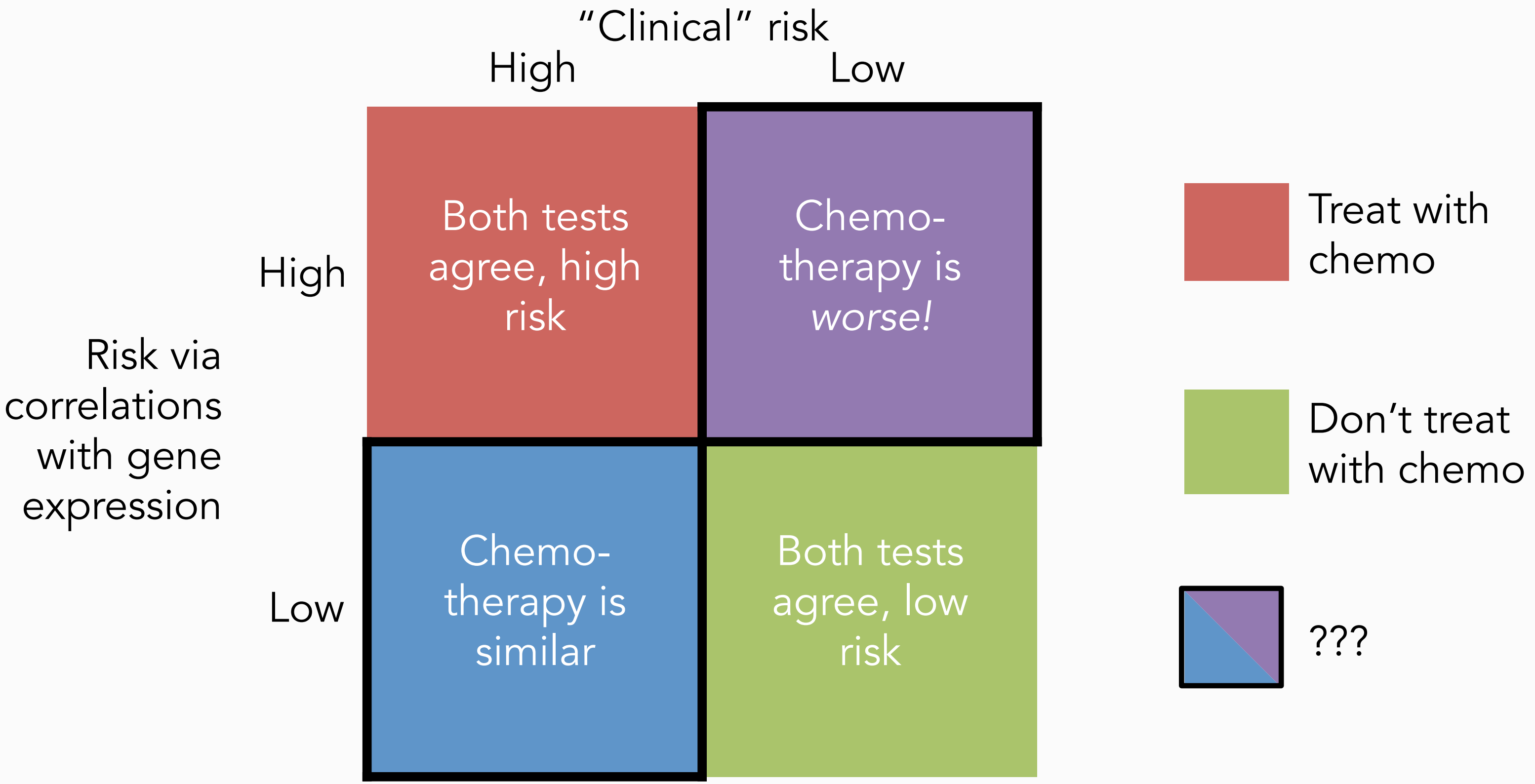
- Fancy paper from 2002 (van't Veer et al.) found 70 genes correlated with developing breast cancer
- Of course the correlations were optimal, post-hoc. But did it generalize?

Real-world testing



(Cardoso et al., 2016)

Real-world testing



(Cardoso et al., 2016)

Real-world testing

Risk via correlations with gene expression	"Clinical" risk: High	"Clinical" risk: Low
High	Both tests agree, high risk	Chemo-therapy is *worse!*
Low	Chemo-therapy is similar	Both tests agree, low risk

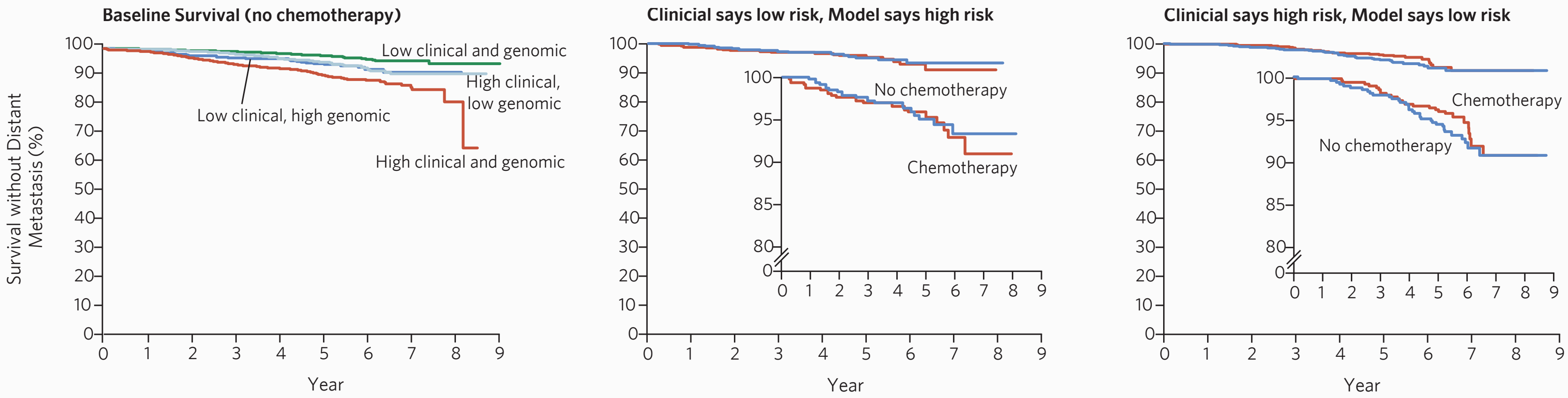
Treat with chemo

Don't treat with chemo

(Still: whose data went into the model? Who were the subjects in the experiment?)

(Cardoso et al., 2016)

Real-world testing: details



- Before experiment (training data)
- High model risk, low clinical risk: randomize. Chemo *worse*!
- Low model risk, high clinical risk: chemo makes no difference

(Cardoso et al., 2016)

Conclusion

Final points

- To use prior data is potentially to "optimize to the status quo" (Carr, 2014)
- Every methodology has limitations, with consequences
- No methodology is *inherently* more oppressive than others; connections to power are what make it so

Other issues

- Performativity: Idea that models, when applied, "reformat and reorganize the phenomena the models purport to describe" (Healy, 2015)
- Quantification can't get at meaning-making (Patton, 2015)
- Statistics (and machine learning) assumes the world is entities with fixed properties; "it is striking how absolutely these assumptions contradict those of the major theoretical traditions of sociology" (Abbott, 1988)
- Only valuing quantitative forms of evidence is a tool to deny lived experience (Lanius, 2011; Benjamin, 2019)
- No amount of data is ever "enough" (Harford, 2014)

References (1/4)

Abbott, Andrew. "Transcending General Linear Reality." *Sociological Theory* 6, no. 2 (1988): 169–186. https://dx.doi.org/10.2307/202114.

Benjamin, Ruha. *Race After Technology: Abolitionist Tools for the New Jim Code*. Cambridge, UK: Polity Press, 2019.

Boden, Margaret A. *Mind as Machine: A History of Cognitive Science*. Oxford: Oxford University Press, 2006.

Bouk, Dan. *How Our Days Became Numbered: Risk and the Rise of the Statistical Individual*. Chicago: The University of Chicago Press, 2015.

Breiman, Leo. "Statistical Modeling: The Two Cultures." *Statistical Science* 16, no. 3 (2001): 199–231. https://dx.doi.org/10.1214/ss/1009213726

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

Cardoso, Fatima, 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, Theodora Goulioti, Susan Knox, Erika Matos, Bart Meulemans, Peter A. Neijenhuis, Ulrike Nitz, Rodolfo Passalacqua, Peter Ravdin, Isabel T. Rubio, Mahasti Saghatchian, Tineke J. Smilde, Christos Sotiriou, Lisette Stork, Carolyn Straehle, Geraldine Thomas, Alastair M. Thompson, Jacobus M. van der Hoeven, Peter Vuylsteke, René Bernards, Konstantinos Tryfonidis, Emiel Rutgers, and Martine Piccart. 2016. "70-Gene Signature as an Aid to Treatment Decisions in Early-Stage Breast Cancer." *New England Journal of Medicine* 375, no. 8 (2016), 717–729. https://dx.doi.org/10.1056/NEJMoa1602253.

Carr, Nicholas. "The Limits of Social Engineering." *MIT Technology Review*, May/June 2014. https://www.technologyreview.com/s/526561/the-limits-of-social-engineering/.

References (2/4)

Despois, Julien. "Adversarial Examples and their implications - Deep Learning bits #3." *Hacker Noon*, 11 Aug 2017. https://hackernoon.com/the-implications-of-adversarial-examples-deep-learning-bits-3-4086108287c7.

Eykholt, Kevin, Ivan Evtimov, Earlence Fernandes, Bo Li, Amir Rahmati, Chaowei Xiao, Atul Prakash, Tadayoshi Kohno, and Dawn Song. "Robust Physical-World Attacks on Deep Learning Visual Classification." *CVPR*, 2018.

Fergus, Devin. 2013. "The Ghetto Tax: Auto Insurance, Postal Code Profiling, and the Hidden History of Wealth Transfer." In *Beyond Discrimination: Racial Inequality in a Postracial Era*, edited by Fredrick C. Harris and Robert C. Lieberman, 277–316. New York: Russel Sage Foundation, 2013. https://aaas.osu.edu/sites/aaas.osu.edu/files/DFergus_Ghetto%20Tax-1%20Auto%20Insurance%20Prop103.pdf.

Fisher, R. A. "On the Mathematical Foundations of Theoretical Statistics." *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* 222 (1922): 309–368. https://dx.doi.org/10.1098/rsta.1922.0009.

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

Gray, Mary L. and Siddharth Suri. *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass*. Boston: Houghton Mifflin Harcourt, 2019.

Harford, Tim. "Big Data: Are We Making a Big Mistake?" *Significance*, December 2014, 14–29. https://dx.doi.org/10.1111/j.1740-9713.2014.00778.x.

Healy, Kieran. "The Performativity of Networks." *European Journal of Sociology* 56, no. 2 (2015): 175–205. https://dx.doi.org/10.1017/S0003975615000107.

Horan, Caley. *Actuarial Age: Insurance and the Emergence of Neoliberalism in the Postwar United States*. PhD diss., University of Minnesota, 2011.

References (3/4)

Jones, Matthew L. 2018. "How We Became Instrumentalists (Again): Data Positivism since World War II." *Historical Studies in the Natural Sciences* 48, no. 5 (2018), 673–684. https://dx.doi.org/10.1525/hsns.2018.48.5.673.

Lanius, Candice. "Fact Check: Your Demand for Statistical Proof is Racist." *Cyborgology* blog, January 15, 2015. https://thesocietypages.org/cyborgology/2015/01/12/fact-check-your-demand-for-statistical-proof-is-racist/.

Lazer, David, Ryan Kennedy, Gary King, and Alessandro Vespignani. "The Parable of Google Flu: Traps in Big Data Analysis." *Science* 343, no. 6176 (2014): 1203–1205. https://dx.doi.org/10.1126/science.1248506.

Messerli, Franz H. "Chocolate Consumption, Cognitive Function, and Nobel Laureates." *The New England Journal of Medicine*, 367 (2012): 1562–1564. https://dx.doi.org/10.1056/NEJMon1211064.

Mitchell, Tom M. *Machine Learning*. McGraw Hill, 1997.

Norvig, Peter. "On Chomsky." 2010. http://norvig.com/chomsky.html.

Patton, Michael Quinn. "The Nature, Niche, Value, and Fruit of Qualitative Inquiry." In *Qualitative Research & Evaluation Methods: Integrating Theory and Practice*, 4th edition, 2–44. SAGE Publications, Inc., 2014. https://uk.sagepub.com/sites/default/files/upm-binaries/64990_Patton_Ch_01.pdf.

Rescher, Nicholas. *Predicting the Future: An Introduction to the Theory of Forecasting*. Albany, New York: State University of New York Press, 1998.

Rose, Todd. *The End of Average: How We Succeed in a World That Values Sameness*. New York: HarperOne, 2016. See excerpt at https://www.thestar.com/news/insight/2016/01/16/when-us-air-force-discovered-the-flaw-of-averages.html. Animated video: https://vimeo.com/237632676.

References (4/4)

Santillana, Mauricio, Wendong Zhang, Benjamin M. Althouse, and John W. Ayers. "What Can Digital Disease Detection Learn from (an External Revision to) Google Flu Trends?" *American Journal of Preventive Medicine* 47, no. 3 (2014): 341–347. http://dx.doi.org/10.1016/j.amepre.2014.05.020.

van't Veer, Laura J., Hongyue Dai, Marc J. van de Vijver, Yudong D. He, Augustinus A. M. Hart, Mao Mao, Hans L. Peterse, Karin van der Kooy, Matthew J. Marton, Anke T. Witteveen, George J. Schreiber, Ron M. Kerkhoven, Chris Roberts, Peter S. Linsley, René Bernards, and Stephen H. Friend. 2002. "Gene Expression Profiling Predicts Clinical Outcome of Breast Cancer." *Nature* 415, no. 6871 (2002): 530–536. https://dx.doi.org/10.1038/415530a.

Wasserman, Larry A. "Rise of the Machines." In *Past, Present, and Future of Statistical Science*, 525–536. Boca Raton, FL: Chapman and Hall/CRC, 2013. http://www.stat.cmu.edu/~larry/Wasserman.pdf.