BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

Three Open Problems for Historians of AI

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

Towards a History of Artificial Intelligence
Workshop at Columbia University, 24 May 2019
Slides: https://mominmalik.com/three_problems.pdf

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

Contact me for technical advising!
<momin.malik@gmail.com>

BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

The Three Problems

1. **When did ML become statistics?**
2. **When did algorithms become models?**
3. **What is the genealogy of prediction?**

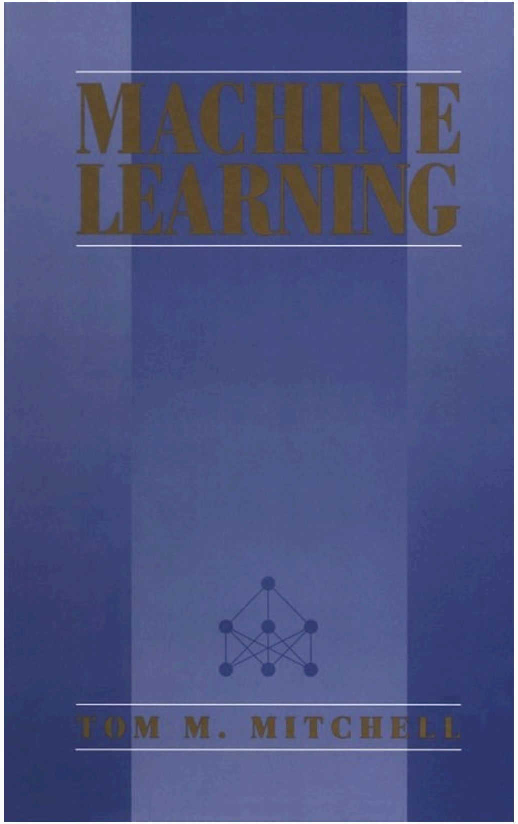
Background and goals

- My interest, and "internalist" mission, is driven by fundamental skepticism
- There are never any discontinuities; but also, disciplinary histories are suspect
- I want rigorous histories to cite for my audiences, collaborators, colleagues

New narratives

1. When did ML become statistics?

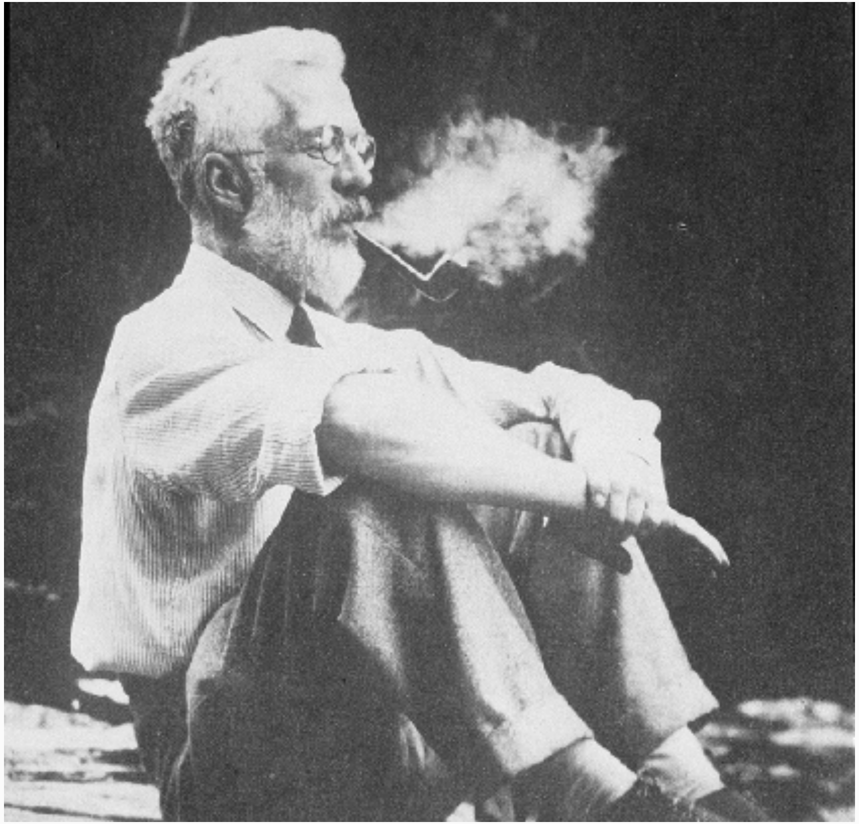
Machine Learning $\not\Rightarrow$ Data/Statistics



- "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*." –Mitchell, *Machine Learning*, 1997

Data/Statistics $\not\Rightarrow$ Probability

"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



BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

Probability comes from *gambling*

- "It is remarkable that a science which began with the **consideration of games of chance** should have become the most important object of human knowledge."
  –Laplace, *Théorie Analytique des Probabilitiés*, 1812

The story I'd *like* to tell...

- ML *started off* trying to make "learning machines"
- That failed
- They found correlations in data could achieve the tasks they were trying to do
- They switched to doing statistics, but called it the same thing
- Not perfectly accurate, but...

Exhibit A: Norvig's 14 years of effort

- "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**". –Norvig, "On Chomsky", 2010

Exhibit B: Boden's take

- "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**." –Boden, *Mind as Machine*, 2006

BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY
AT HARVARD UNIVERSITY

Exhibit C: "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**." –Wasserman, "Rise of the Machines", 2014

BERKMAN KLEIN CENTER FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

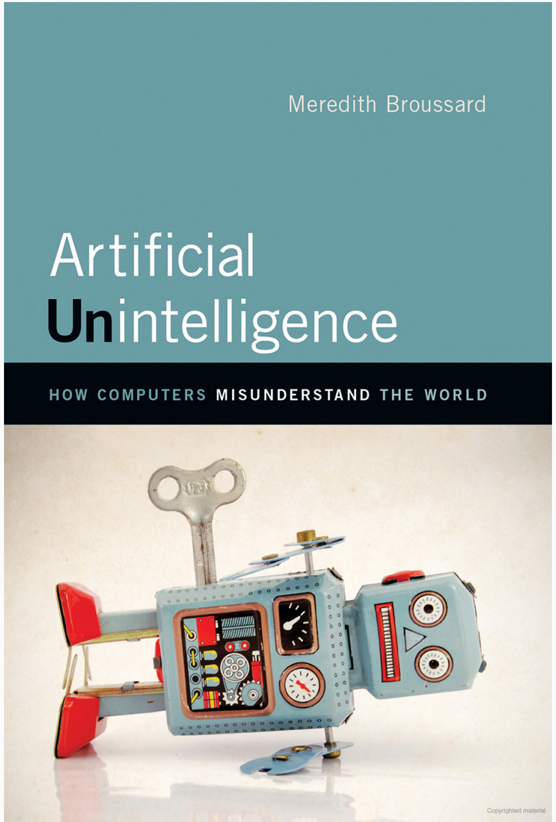
Exhibit D: Breiman's "second" culture

- "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**." –Breiman, "The Two Cultures", 2001

BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

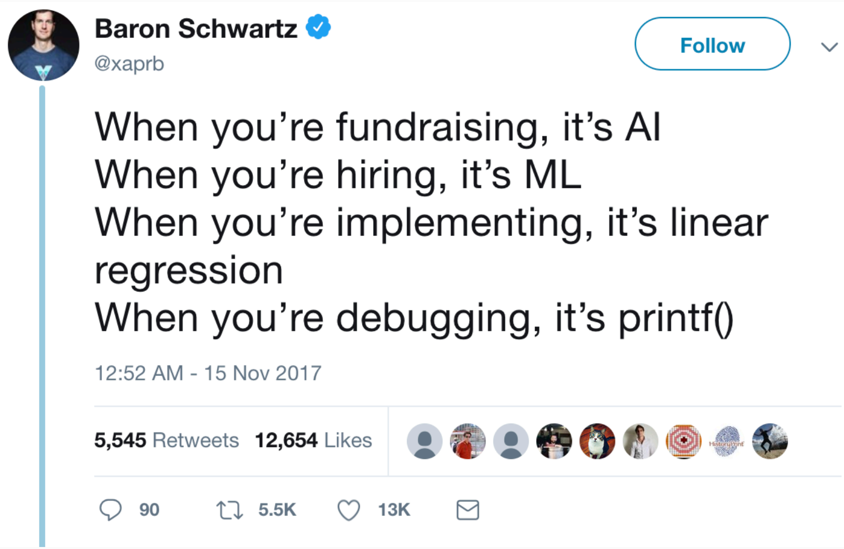
Why does it matter?

Status quo: Useful ignorance



- Story in Ch. 3, "Hello, 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."

Fight this with demystification

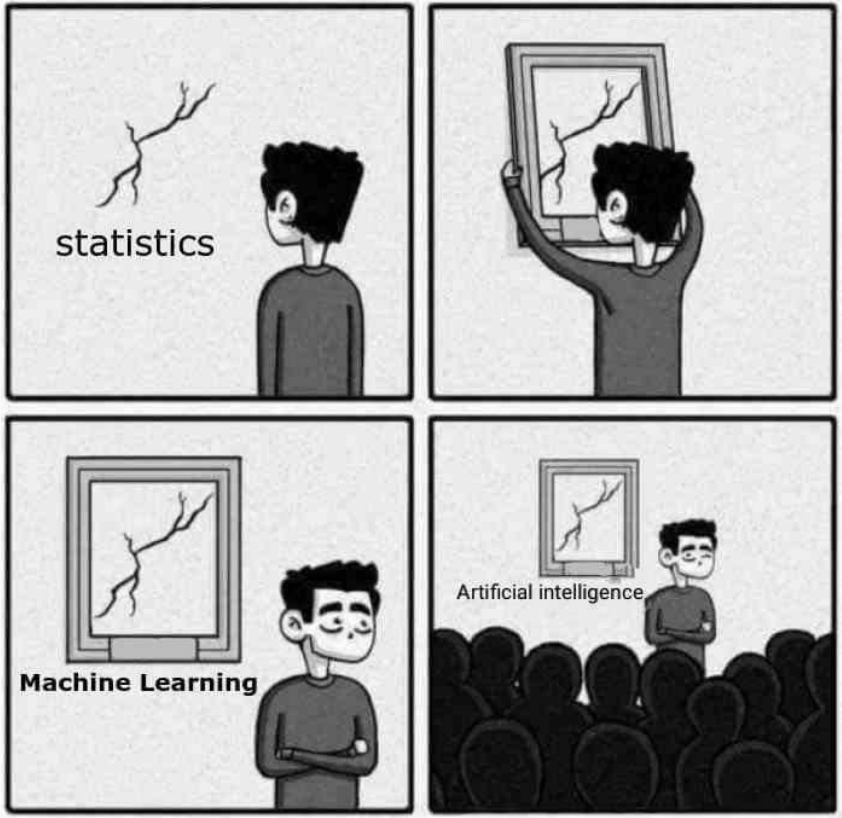

Baron Schwartz
@xaprb
Follow

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

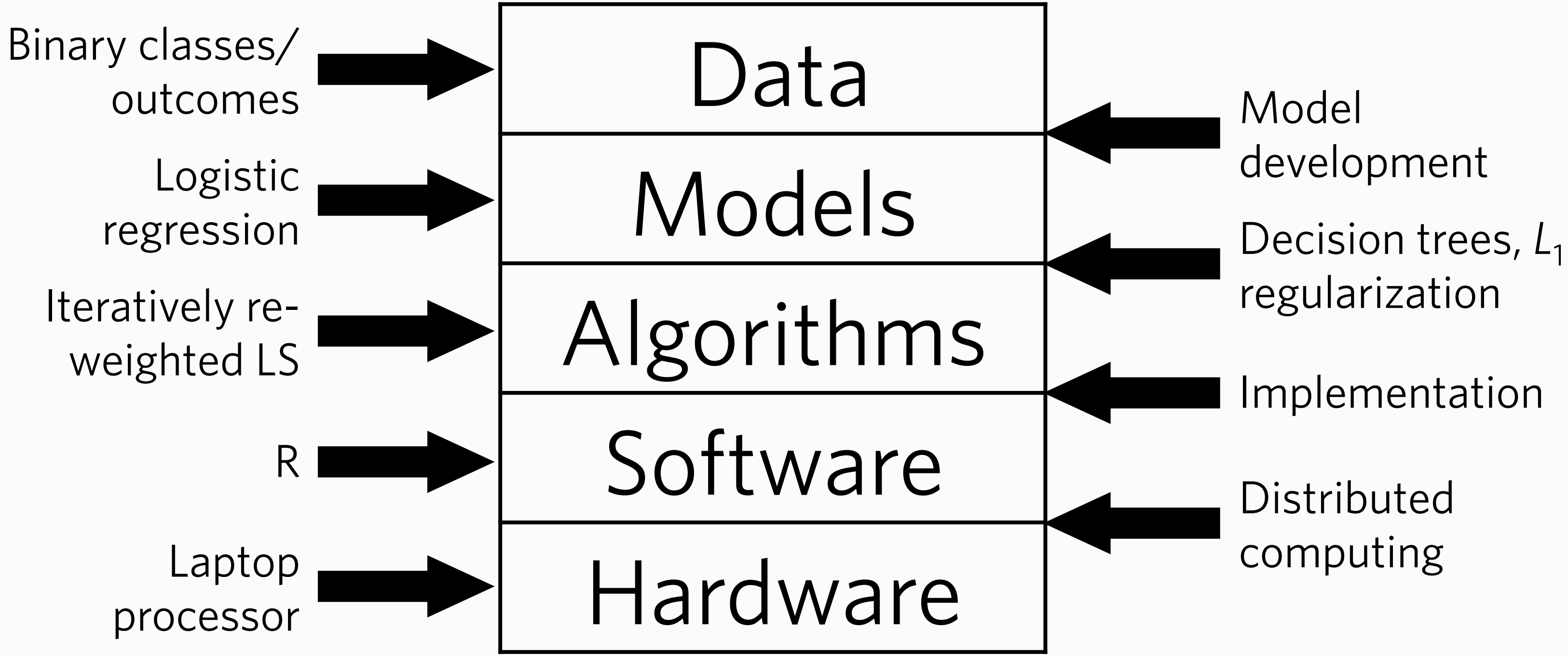


BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

2. When did algorithms become models?

BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

Confusion in *levels of abstraction*



Advantages of rhetorical distancing?

- "neural networkers tend to **ignore the distributional assumptions** they have made, whereas statisticians explore their consequences." –Boden
- Language of "improvement" avoids analysis of convergence?
- Other psychological/political benefits?

Avoids critiques of modeling

Can an Algorithm be Wrong?

How do we know if we are where it's at? **Tarleton Gillespie** explores the controversy over Twitter Trends and the algorithmic 'censorship' of #occupywallstreet.

THROUGHOUT the Occupy Wall Street protests, participants and supporters used Twitter (among other tools) to coordinate, debate, and publicize their efforts. But amidst the enthusiasm a concern surfaced: even as the protests were gaining strength and media coverage, and talk of the movement on Twitter was surging, the term was not "Trending." A simple list of ten terms provided by Twitter on their homepage, Twitter Trends digests the 250 million tweets sent every day and indexes the most vigorously discussed terms at that moment, either globally or for a user's chosen country or city. Yet, even in the cities where protests were happening, including New York, when tweets using the term #occupywallstreet seem to spike, the term did not Trend. Some suggested that Twitter was deliberately dropping the term from its list, and in doing so, preventing it from reaching a wider audience.

The charge of censorship is a revealing one. It suggests, first, that many are deeply invested in the Twitter network as a political tool, and that some worry that Twitter's

cupywallstreet's absence is that Twitter "censored" it implies that Trends is otherwise an accurate barometer of the public discussion. For some, this glitch could only mean deliberate human intervention into what should be a smoothly-running machine. The workings of these algorithms are political, an important terrain upon which political battles about visibility are being fought (Grimmelmann 2009). Much like taking over the privately owned Zuccotti Park in Manhattan in order to stage a public protest, more and more of our online public discourse is taking place on private communication platforms like Twitter. These providers offer complex algorithms to manage, curate, and organize these massive networks. But there is no tension between what we understand these algorithms to be, what we need them to be, and what they in fact are. We do not have a sufficient vocabulary for assessing the intervention of these algorithms. We're not adept at appreciating what it takes to design a tool like Trends – one that appears to effortlessly identify what's going on, yet also makes distinct and motivated choices. We don't have a language for the unexpected associations algorithms make, beyond the intention (or even comprehension) of their designers (Ananny 2011). Most importantly, we have not fully recognized how these algorithms attempt to produce representations of the wants or concerns of the public, and as such, run into the classic problem of political representation: who

UNIVERSITY OF WISCONSIN-MADISON
MATHEMATICS RESEARCH CENTER

ROBUSTNESS IN THE STRATEGY OF SCIENTIFIC MODEL BUILDING†

G. E. P. Box

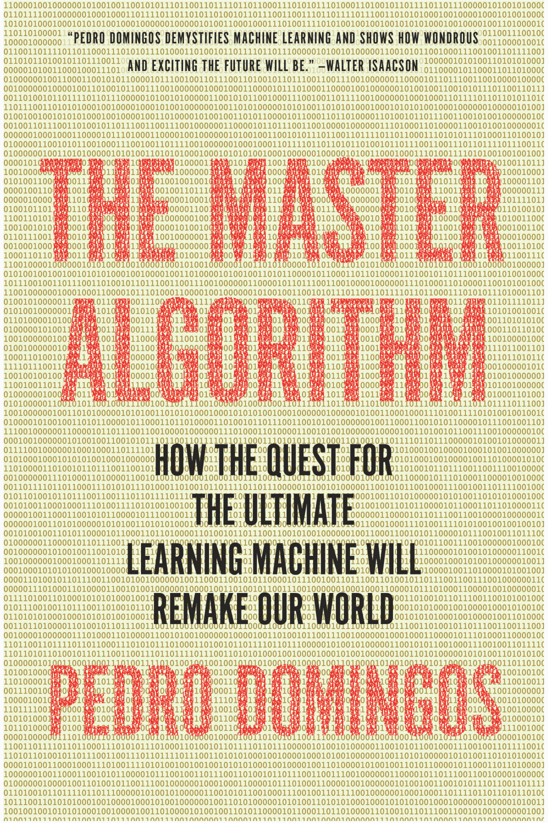
Technical Summary Report #1954
May 1979

ABSTRACT

ALL MODELS ARE WRONG BUT SOME ARE USEFUL

BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

Allows for megalomania



- "Like somebody volunteered to be a straw horse" –Colleague of mine
- Given Box's dictum, would anybody take seriously "The Master Model"?

BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

Why does it matter?

Study the right topic!

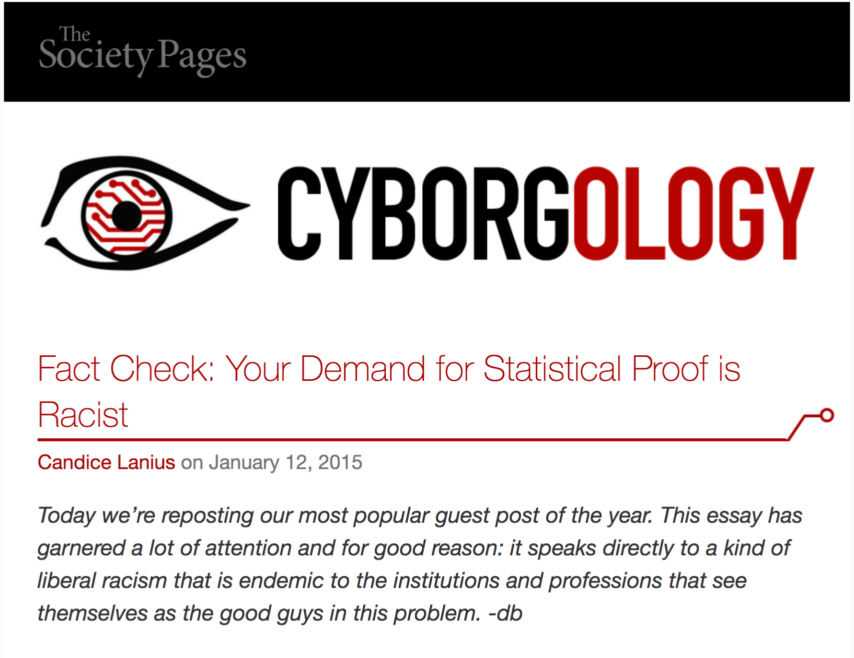
- “I hopped over to the Department of Management at LSE to visit my new friend [and] I asked him: **‘What is an algorithm?’** In other words, what is the scope of things I can expect will get discussed at a conference about algorithms? ‘That’s a very good question,’ Keith replied, ‘When I started, there was a fairly precise meaning to the term. An algorithm was a set of rules, which would generate an optimum answer to the problem that you’d posed it. It was a statement of rules that gave you the best possible answer in a finite amount of time.’” –Poon, “Response to Gillespie”, 2013

Apply critiques of statistics!

- "[Statistical models] **assumes the social world consists of fixed entities (the units of analysis) that have attributes (the variables). These attributes interact, in causal or actual time, to create outcomes, themselves measurable as attributes of the fixed entities… it is striking how absolutely these assumptions contradict those of the major theoretical traditions of sociology**. Symbolic interactionism rejects the assumption of fixed entities and makes the meaning of a given occurrence depend on its location—within an interaction, within an actor's biography, within a sequence of events. Both the Marxian and Weberian traditions deny explicitly that a given property of a social actor has one and only one set of causal implications…" –Abbott, "Transcending General Linear Reality", 1988

BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

Apply critiques of statistics!



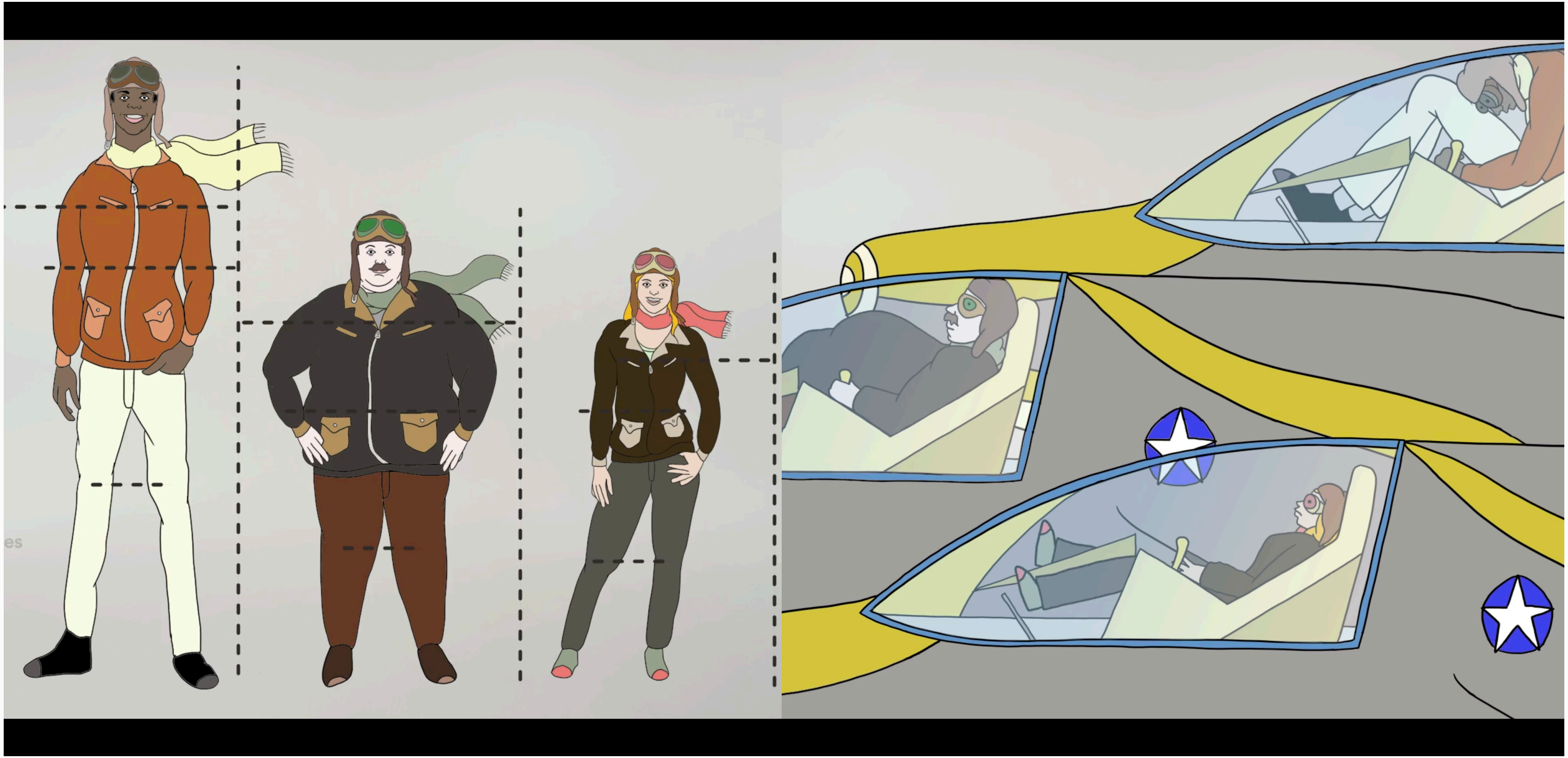
Fact Check: Your Demand for Statistical Proof is Racist

Candice Lanius on January 12, 2015

Today we're reposting our most popular guest post of the year. This essay has garnered a lot of attention and for good reason: it speaks directly to a kind of liberal racism that is endemic to the institutions and professions that see themselves as the good guys in this problem. -db

- "A white woman can say that a neighborhood is 'sketchy' and most people will smile and nod. She felt unsafe, and we automatically trust her opinion. A black man can tell the world that every day he lives in fear of the police, and suddenly everyone demands statistical evidence to prove that his life experience is real."

Apply critiques of statistics!



3. What is the genealogy of “prediction”?

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.

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." –Gayo-Avello, "I Wanted to Predict Elections with Twitter and all I got was this Lousy Paper", 2012

Prediction is post-hoc correlation

- Non-causal correlations can fit the data really well!
- Google Flu Trends: half flu detector, half winter detector
- Where did this come from?

BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

Aspirational naming of AI?

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 ... ble. ... 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.

No, much older: Pearson, 1928

equal to certainty, would be found to be lumped up. Let us suppose for illustration that a mark was made on our scale of skin pigmentation and the number of individuals in a particular race with a darker pigmentation was found to be r, and without knowledge of other races, we wanted to predict something about the percentage of dark individuals in the sampled population. Now clearly with such a method of estimating dark individuals in a race, the frequency distri-

Shannon, 1950

Prediction and Entropy of Printed English

By C. E. SHANNON

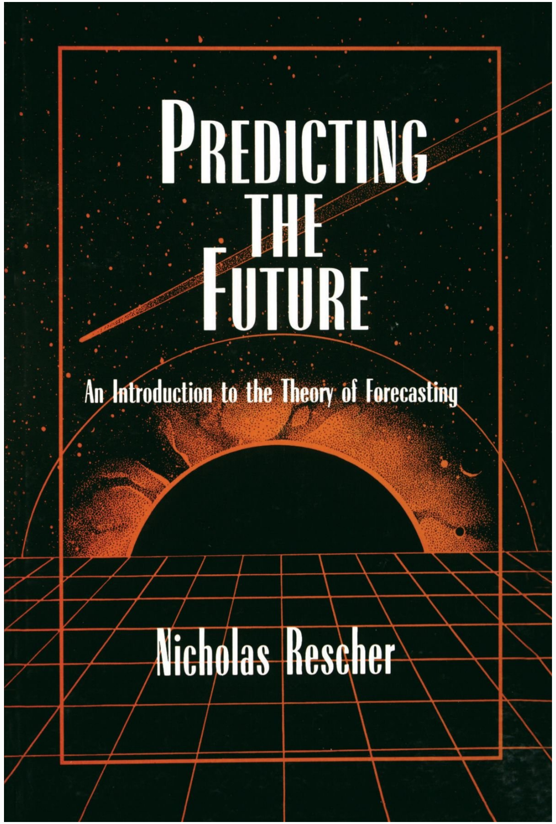
(Manuscript Received Sept. 15, 1950)

A new method of estimating the entropy and redundancy of a language is described. This method exploits the knowledge of the language statistics possessed by those who speak the language, and depends on experimental results in prediction of the next letter when the preceding text is known. Results of experiments in prediction are given, and some properties of an ideal predictor are developed.

Solomonoff, 1956

The inductive inference machine takes categories that have been useful in the past and, by means of a small set of transformations, derives new categories that have reasonable likelihood of being useful in the future. These are then tested empirically for usefulness in prediction and the new useful ones are combined with old useful categories to create newer ones. These, in turn, are tested and the process is repeated again and again.

Accidentally, uniquely dangerous?



- "In all times and places, decision makers have looked to predictive counselors of some sort—putative experts, be they religious or secular, to guide them regarding the auguries of the gods, the stars, or the inexorable decrees of fate or of nature."
- (Or perhaps deeply revealing?)

Conclusion

Summary

- Certain historical work would be enormously helpful for my political project:
- Show that ML became statistics/statistical
- Show algorithms became *models*, and *modeling* logic is what we worry about
- Show where "prediction" comes from, to show that it's not actually prediction

https://flic.kr/p/SQVdB4

My own prediction

- The hype around AI and machine learning will die in 5-10 years as it fails to achieve new successes
- *Modeling* will remain a persistent issue for the remainder of human civilization
- AI might be an interesting historical case study of aspiration/arrogance, but literature about modeling will retain contemporary relevance forever

References

"What Will Become of Us?", *New York Times Magazine* (The Tech & Design issue), November 14, 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.

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

Asur, Sitaram 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.

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

Box, George E. P. "Robustness in the Strategy of Scientific Model Building." MRC Technical Summary Report #1954. University of Wisconsin-Madison Mathematics Research Center, March 29, 1976.

Breiman, Leo. "Statistical Modeling: The Two Cultures (with Comments and a Rejoinder by the Author)." *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*. MIT Press, 2018.

Domingos, Pedro. The *Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*. Basic Books, 2015.

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." arXiv, 28 April 2012. https://arxiv.org/abs/1204.6441.

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

Gillespie, Tarleton. "Can An Algorithm Be Wrong?" *limn* 2 (2012 March): 17–20. https://limn.it/articles/can-an-algorithm-be-wrong/.

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://www.jstor.org/stable/91208.

Jones, Matthew L. "How We Became Instrumentalists (Again): Data Positivism since World War II." *Historical Studies in the Natural Sciences* 48, no. 5 (2018): 673–684. http://www.columbia.edu/~mj340/H SNS4805 12 Jones.pdf.

Lanius, Candice. "Fact Check: Your Demand for Statistical Proof is Racist." *Cyborgology* blog, January 15, 2015. https://thesocietypages.org/cyborgolo gy/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 Trends: Traps in Big Data Analysis." *Science* 343 (14 March 2014): 1203–1205. https://dx.doi.org/10.1126/science.1248506.

Lipton, Zachary C. "The Myth of Model Interpretability." *KDnuggets* 15, no. 13 (April 2015). https://www.kdnuggets.com/2015/04 /model-interpretability-neural-networks-deep-learning.html; Lipton, Zachary C. "The Mythos of Model Interpretability." *ACM Queue* 16, no. 3 (June 2018): 31–57, https://dx.doi.org/10.1145/3236386.3241340.

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

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

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

Norvig, Peter. "On Chomsky and the Two Cultures of Statistical Learning." 2010. http://norvig.com/chomsky.html.

Obermeyer, Ziad 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.

Pearson, Karl. "On a Method of Ascertaining Limits to the Actual Number of Marked Members in a Population of Given Size from a Sample." *Biometrika* 20A, no. 1-2 (1928, July): 149-174. https://www.jstor.org/stable/2331944.

Poon, Martha. "Response to Tarleton Gillespie's 'The Relevance of Algorithms'." *Governing Algorithms* Conference, NYU, May 17, 2013.

Rescher, Nicholas. *Predicting the Future: An Introduction to the Theory of Forecasting*. SUNY Press, 1997.

Rose, Todd. *End of Average: How We Succeed in a World That Values Sameness*. HarperOne, 2016. Images: Screengrabs from Future for Learning, "Jet Cockpits and The End of Average." https://vimeo.com/237632676.

Shannon, Claude E. "Prediction and Entropy of Printed English." *The Bell System Technical Journal* 30, no. 1 (1951, January): 50–64. https://dx.doi.org/10.1002/j.1538-7305.1951.tb01366.x.

Solomonoff, Ray J. "An Inductive Inference Machine." Dartmouth College, NH, August 14, 1956.

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