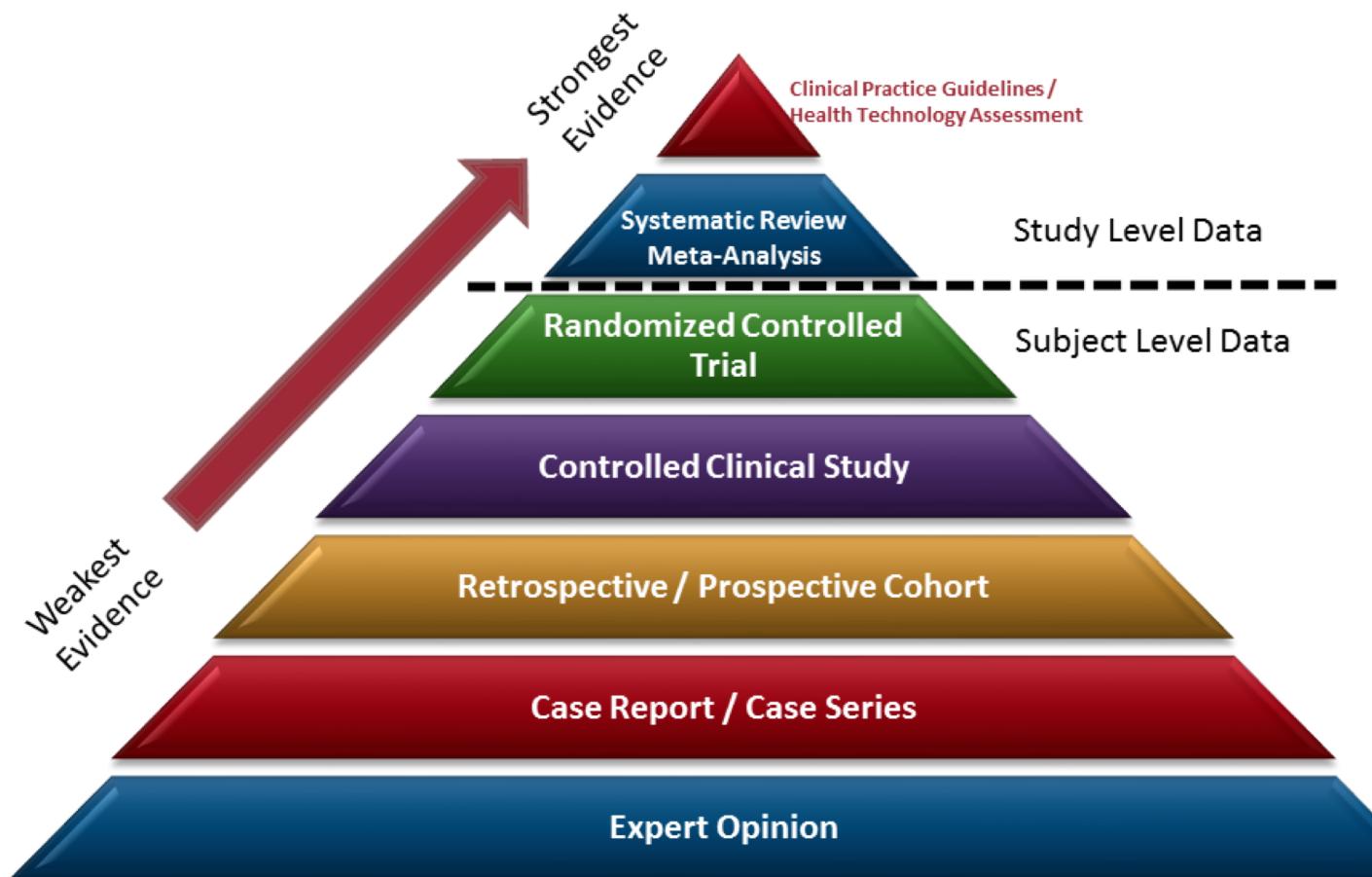


Evidence-based decision making

Rui Mata, FS 2021



https://en.wikipedia.org/wiki/Levels_of_evidence

Policy capturing in practice



<https://evidencebaseddm.formr.org>

1

Choosing Prediction Over Explanation in Psychology: Lessons From Machine Learning

Tal Yarkoni and Jacob Westfall

University of Texas at Austin

Perspectives on Psychological Science
2017, Vol. 12(6) 1100–1122
© The Author(s) 2017
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/1745691617693395
www.psychologicalscience.org/PPS


Abstract

Psychology has historically been concerned, first and foremost, with explaining the causal mechanisms that give rise to behavior. Randomized, tightly controlled experiments are enshrined as the gold standard of psychological research, and there are endless investigations of the various mediating and moderating variables that govern various behaviors. We argue that psychology's near-total focus on explaining the causes of behavior has led much of the field to be populated by research programs that provide intricate theories of psychological mechanism but that have little (or unknown) ability to predict future behaviors with any appreciable accuracy. We propose that principles and techniques from the field of machine learning can help psychology become a more predictive science. We review some of the fundamental concepts and tools of machine learning and point out examples where these concepts have been used to conduct interesting and important psychological research that focuses on predictive research questions. We suggest that an increased focus on prediction, rather than explanation, can ultimately lead us to greater understanding of behavior.

2

The era of blind faith in big data must end | Cathy O'Neil

 Watch later  Share



Yarkoni & Westfall

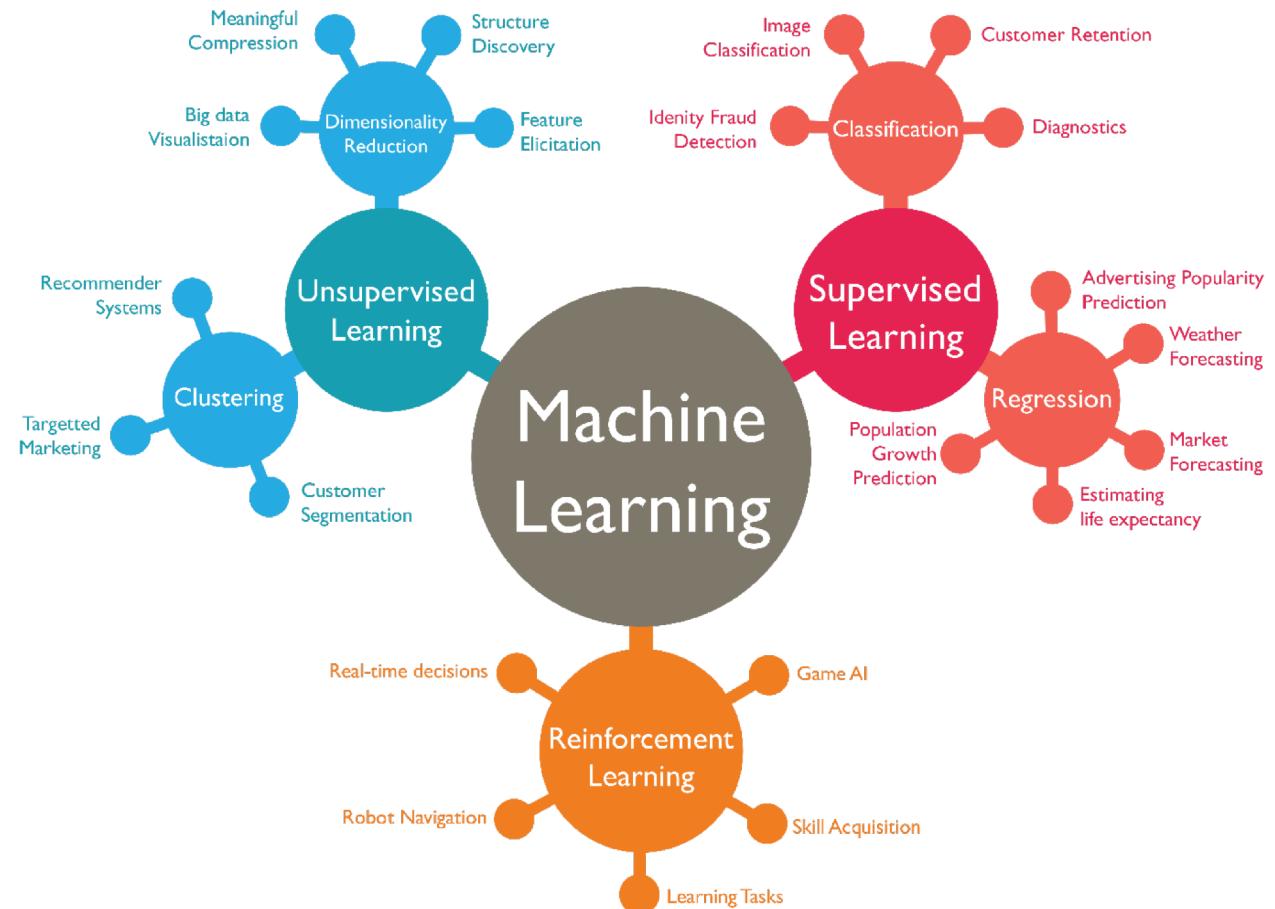
“research programs that provide intricate theories of psychological mechanism but that have little (or unknown) ability to predict future behaviors”

O’Neil

“.. algorithms are opinions embedded in code”!

Actuarial judgment today = machine learning...

“Machine learning is a discipline focused on two interrelated questions: How can one construct computer systems that automatically improve through experience? and What are the fundamental statistical-computational-information-theoretic laws that govern all learning systems, including computers, humans, and organizations?”



Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255–260. <http://doi.org/10.1126/science.aaa8415>

Limitations of actuarial judgment: Overfitting

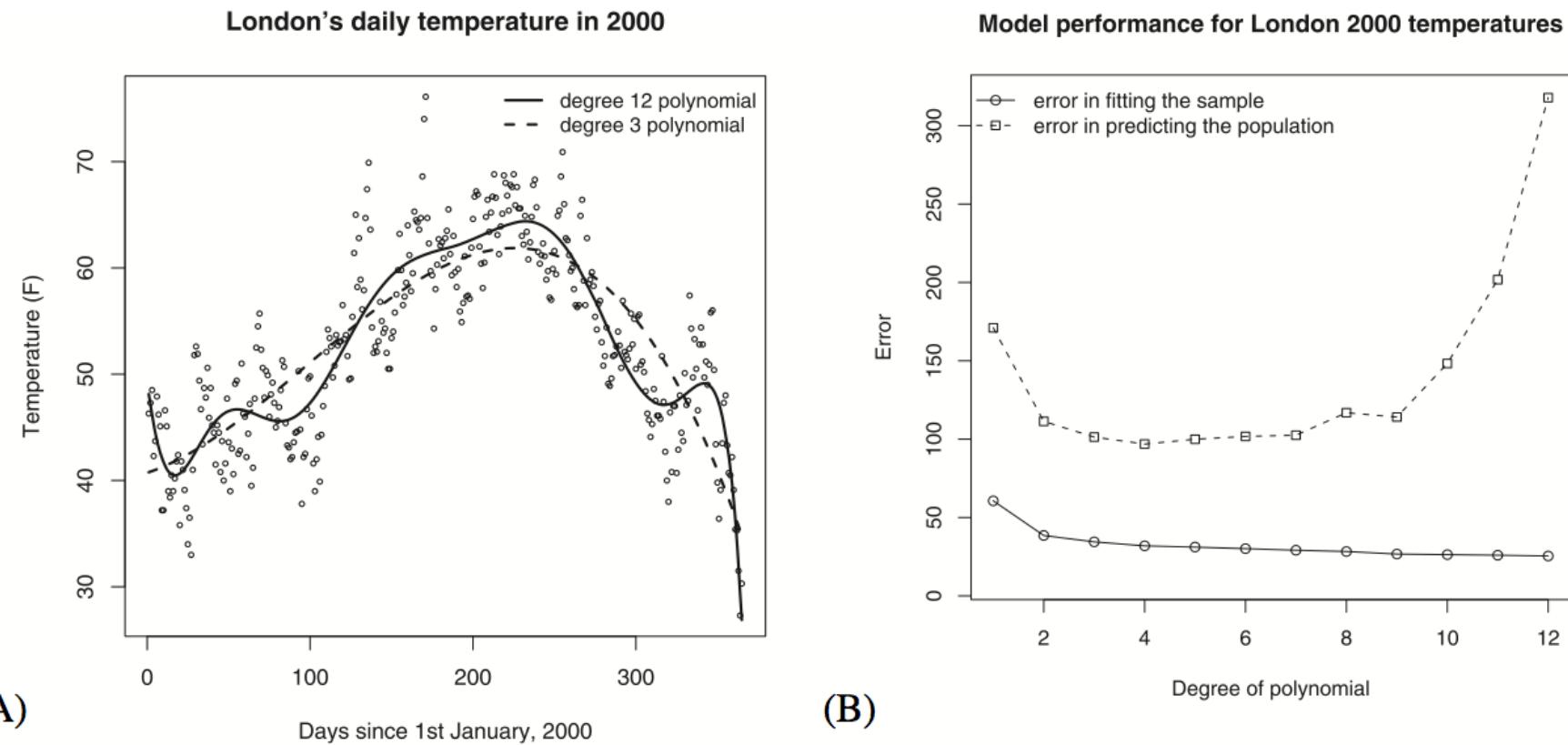


Fig. 3. Plot (A) shows London's mean daily temperature in 2000, along with two polynomial models fitted with using the least squares method. The first is a degree-3 polynomial, and the second is a degree-12 polynomial. Plot (B) shows the mean error in fitting samples of 30 observations and the mean prediction error of the same models, both as a function of degree of polynomial.

Gigerenzer, G., & Brighton, H. (2009). Homo Heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science*, 1(1), 107–143. doi:10.1111/j.1756-8765.2008.01006.x

How can one avoid overfitting?

- **Regularization:** Use regularisation in the estimation of model parameters (e.g. ridge or lasso regression)

$$\text{Regularized loss} = \sum_i^n (y_i - \hat{y}_i)^2 + \lambda \sum_j^p f(\beta_j))$$

lasso regression: $|\beta|$ (β s are reduced in size, resulting in automatic feature selection, with some β s becoming zero)

ridge regression: β^2 (squaring reduces the size of extreme β s).

elastic net: $|\beta| + \beta^2$ (the best of both worlds)

How can one avoid overfitting?

- **Cross-validation:** Compare models in how well they predict out-of-sample (cross-validation/prediction)

Training

Test (holdout)

Yarkoni, T. & Westfall, J. (2017). Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on Psychological Science*, 12(6), 1100-1122.

Complex algorithms may lead to only small increases in performance!

Table 1

Performance of Different Algorithms in Predicting House Values

Method	Prediction performance (R^2)		Relative improvement over ordinary least squares by quintile of house value				
	Training sample	Hold-out sample	1st	2nd	3rd	4th	5th
Ordinary least squares	47.3%	41.7% [39.7%, 43.7%]	-	-	-	-	-
Regression tree tuned by depth	39.6%	34.5% [32.6%, 36.5%]	-11.5%	10.8%	6.4%	-14.6%	-31.8%
LASSO	46.0%	43.3% [41.5%, 45.2%]	1.3%	11.9%	13.1%	10.1%	-1.9%
Random forest	85.1%	45.5% [43.6%, 47.5%]	3.5%	23.6%	27.0%	17.8%	-0.5%
Ensemble	80.4%	45.9% [44.0%, 47.9%]	4.5%	16.0%	17.9%	14.2%	7.6%

Note: The dependent variable is the log-dollar house value of owner-occupied units in the 2011 American Housing Survey from 150 covariates including unit characteristics and quality measures. All algorithms are fitted on the same, randomly drawn training sample of 10,000 units and evaluated on the 41,808 remaining held-out units. The numbers in brackets in the hold-out sample column are 95 percent bootstrap confidence intervals for hold-out prediction performance, and represent measurement variation for a fixed prediction function. For this illustration, we do not use sampling weights. Details are provided in the online Appendix at <http://e-jep.org>.

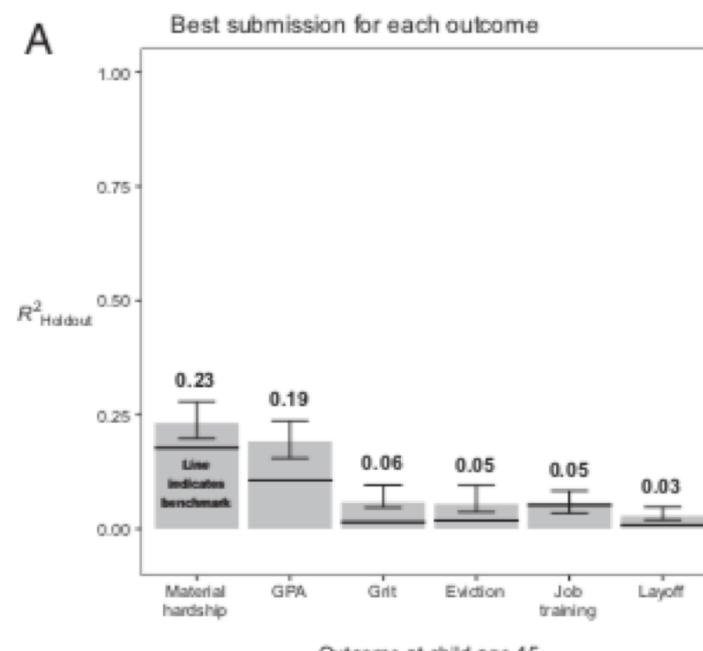
Mullainathan, S., & Spiess, J. (2017). Machine Learning: An applied econometric approach. *Journal of Economic Perspectives*, 31(2), 87–106. <http://doi.org/10.1257/jep.31.2.87>

Complex algorithms do not guarantee predictability!

“How predictable are life trajectories? We investigated this question with a scientific mass collaboration using the common task method; 160 teams built predictive models for six life outcomes using data from the Fragile Families and Child Wellbeing Study, a high-quality birth cohort study. Despite using a rich dataset and applying machine-learning methods optimized for prediction, the best predictions were not very accurate and were only slightly better than those from a simple benchmark model. Within each outcome, prediction error was strongly associated with the family being predicted and weakly associated with the technique used to generate the prediction. Overall, these results suggest practical limits to the predictability of life outcomes in some settings and illustrate the value of mass collaborations in the social sciences.”



Fig. 2. Datasets in the Fragile Families Challenge. During the Fragile Families Challenge, participants used the background data (measured from child's birth to age 9 y) and the training data (measured at child age 15 y) to predict the holdout data as accurately as possible. While the Fragile Families Challenge was underway, participants could assess the accuracy of their predictions in the leaderboard data. At the end of the Fragile Families Challenge, we assessed the accuracy of the predictions in the holdout data.



Salganik, M. J., et al. 2020. (2020). Measuring the predictability of life outcomes with a scientific mass collaboration. PNAS, <http://doi.org/10.7910/DVN/CXSECU>

1

Choosing Prediction Over Explanation in Psychology: Lessons From Machine Learning

Tal Yarkoni and Jacob Westfall
University of Texas at Austin

Perspectives on Psychological Science
2017, Vol. 12(6) 1100–1122
© The Author(s) 2017
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: [10.1177/1346561317709395](https://doi.org/10.1177/1346561317709395)
www.psychologicalscience.org/PPS
SAGE

Abstract

Psychology has historically been concerned, first and foremost, with explaining the causal mechanisms that give rise to behavior. Randomized, tightly controlled experiments are enshrined as the gold standard of psychological research, and there are endless investigations of the various mediating and moderating variables that govern various behaviors. We argue that psychology's near-total focus on explaining the causes of behavior has led much of the field to be populated by research programs that provide intricate theories of psychological mechanism but that have little (or unknown) ability to predict future behaviors with any appreciable accuracy. We propose that principles and techniques from the field of machine learning can help psychology become a more predictive science. We review some of the fundamental concepts and tools of machine learning and point out examples where these concepts have been used to conduct interesting and important psychological research that focuses on predictive research questions. We suggest that an increased focus on prediction, rather than explanation, can ultimately lead us to greater understanding of behavior.

2

The era of blind faith in big data must end | Cathy O'Neil

Watch later Share



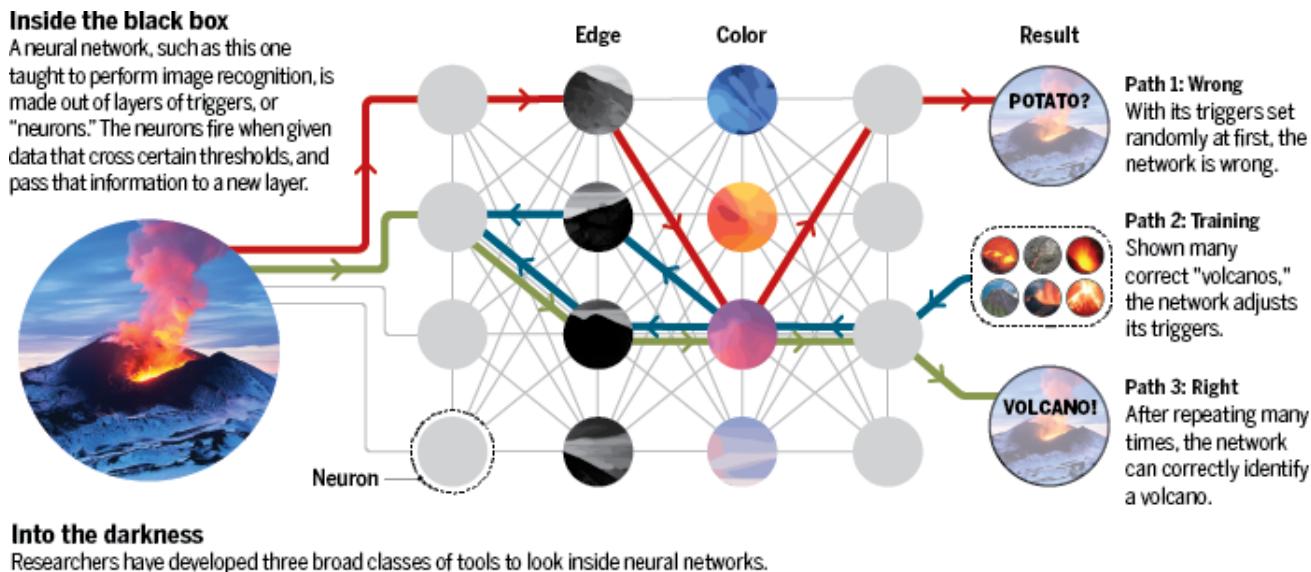
Yarkoni & Westfall

“research programs that provide intricate theories of psychological mechanism but that have little (or unknown) ability to predict future behaviors”

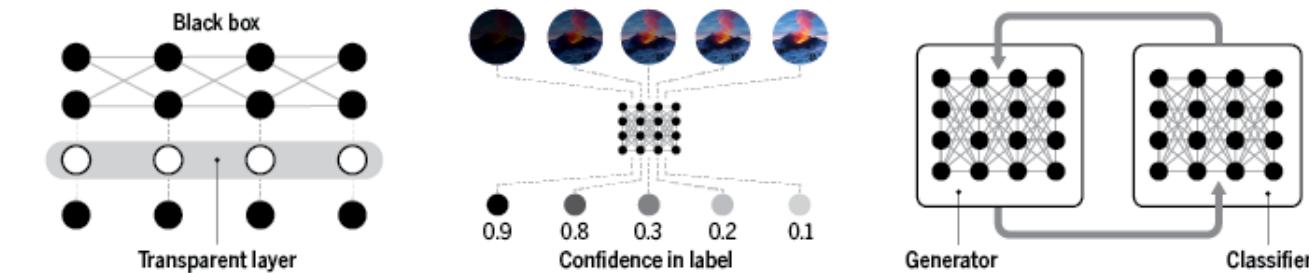
O’Neil

“.. algorithms are opinions embedded in code”!

The need for Explainable AI



Into the darkness
Researchers have developed three broad classes of tools to look inside neural networks.



Voosen (2017) How AI detectives are cracking open the black box of deep learning, *Science*.

See also <https://distill.pub/2018/building-blocks/>



Prof. Dr. iur. Sabine Gless

<p>06 APR 2021 - 04 MAY 2021</p>	<p>18:15 - 19:45</p> <p>Ort: Digitale Veranstaltung</p> <p>Organisator: Doktoratsprogramm "Recht im Wandel"</p>	<p>Ringvorlesung: "Künstliche Intelligenz im Recht"</p> <p>Die Vorlesungsreihe findet jeweils dienstags von 18.15 Uhr bis 19.45 Uhr online statt. Registrierung jetzt möglich!</p> <p>Ab sofort besteht die Möglichkeit sich für die Ringvorlesung (Webinar) unter folgendem Link zu registrieren. Nach der Registrierung erhalten Sie eine Bestätigungs-E-Mail mit Informationen über die Teilnahme am Webinar.</p> <p>> Registrierung</p> <p>Programm </p> <p>Veranstaltung übernehmen als >iCal</p>
------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Ringvorlesung

Algorithm aversion vs. Algorithm Appreciation

A systematic review of algorithm aversion in augmented decision making

Jason W. Burton¹  | Mari-Klara Stein² | Tina Blegind Jensen²

¹Department of Psychological Sciences,
Birkbeck, University of London, London, UK

²Department of Digitalization, Copenhagen
Business School, Frederiksberg, Denmark

Correspondence

Jason W. Burton, Department of Psychological Sciences, Birkbeck, University of London, Malet Street, London, WC1E 7HX, UK.
Email: jasonwilliamburton@gmail.com

Abstract

Despite abundant literature theorizing societal implications of algorithmic decision making, relatively little is known about the conditions that lead to the acceptance or rejection of algorithmically generated insights by individual users of decision aids. More specifically, recent findings of algorithm aversion—the reluctance of human forecasters to use superior but imperfect algorithms—raise questions about whether joint human-algorithm decision making is feasible in practice. In this paper, we systematically review the topic of algorithm aversion as it appears in 61 peer-reviewed articles between 1950 and 2018 and follow its conceptual trail across disciplines. We categorize and report on the proposed causes and solutions of algorithm aversion in five themes: expectations and expertise, decision autonomy, incentivization, cognitive compatibility, and divergent rationalities. Although each of the presented themes addresses distinct features of an algorithmic decision aid, human users of the decision aid, and/or the decision making environment, apparent interdependencies are highlighted. We conclude that resolving algorithm aversion requires an updated research program with an emphasis on theory integration. We provide a number of empirical questions that can be immediately carried forth by the behavioral decision making community.

KEY WORDS

algorithm aversion, augmented decision making, human-algorithm interaction, systematic review

Burton, J. W., Stein, M. K., & Jensen, T. B. (2019). A systematic review of algorithm aversion in augmented decision making. *Journal of Behavioral Decision Making*, 27(11), 1309–20. <http://doi.org/10.1002/bdm.2155>

Algorithm aversion vs. Algorithm Appreciation

False expectations -> decision makers may have incorrect/unrealistic beliefs about the performance of algorithms

Lack of decision control -> decision makers may strive for autonomy

Lack of incentivization -> incentives for algorithmic use may be unclear or misaligned (effort vs. performance)

Combatting intuition -> decision makers may have incorrect (overconfident) beliefs about own intuition

Conflicting concepts of rationality -> lack of a match between algorithm's knowledge and those of the individual (risk vs. uncertainty)

Burton, J. W., Stein, M. K., & Jensen, T. B. (2019). A systematic review of algorithm aversion in augmented decision making. *Journal of Behavioral Decision Making*, 27(11), 1309–20. <http://doi.org/10.1002/bdm.2155>

Summary

- Potential limitations of actuarial methods: overfitting, complexity, lack of interpretability;
- Potential remedies: regularization, cross-validation, transparency/interpretable AI
- Algorithmic aversion vs. algorithmic appreciation: academic debate has centered around reasons for lack of adoption of algorithms in professional/forecasting settings; multifactorial issue (role of training, beliefs in intuition, incentives, etc.).
More interdisciplinary work is needed...