university of groningen

NWO
Netherlands Organisation for Scientific Research

Howson Memorial Conference
LSE 2022

Howson on Induction

With applications to machine learning

★

Jan-Willem Romeijn
University of Groningen

Probability and the logic of induction

Howson’s views on induction and probability were a truly crucial influence on me since I started my PhD.

> *We have solved Hume’s Problem in about the only way it could be solved, by divorcing the justification for inductive reasoning from a justification of its consequences. Inductive reasoning is justified to the extent that it is sound, given appropriate premises.*
>
> Colin Howson, Hume’s Problem, 2000.

In this talk I will illustrate how this view has steered my work and retained its relevance.

Bacon’s epistemo-entomology

The theme of this talk is nicely captured in another influential work on induction, Bacon’s Novum Organon:

> *[Scientists] have been either empirics or dogmatical. The former, like ants, only heap up and use their store, the latter like spiders spin out their own webs. The bee, a mean between both, extracts matter from the flowers of the garden and the field, but works and fashions it by its own efforts.*
>
> Francis Bacon, The New Organon [Book One], 1620.

Machine learning seem like the work of ants. It focuses on collecting data and “letting those data speak for themselves”.

Bees, not ants

As most machine learning experts will tell you, this popular idea of machine learning is mistaken.

The general inevitability of inductive bias is well-known. But it is still a challenge to identify it in concrete cases.

Plan of talk

1. Machine learning in science
2. Concerns over reliability
3. Data-driven psychopathology
4. Learning from a fruit machine
5. Uncovering inductive assumptions
6. Automated text allocation
7. The epistemology of data science

1 Machine learning in science

There are many examples of data-driven methods in the sciences, for prediction and automated model construction:

- Psychiatrists use hierarchical clustering to come up with subtypes of heterogeneous diseases like depression.
- Biomedical researchers employ methods of automated causal discovery to identify mechanisms of gene expression in the cell.
- Linguists employ latent Dirichlet analysis to disclose a corpus of texts and allocate them to thematic clusters.

In these cases the impact of theoretical starting points is difficult to trace. Does that matter?

Methodological concerns

The different nature of the new methods puts the continuity with existing theory under pressure.

And the new methods are “black-boxed”. This makes it hard to hold machine-learning research accountable and motivate policy with it.

Transparency

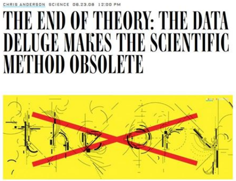
Continuity and accountability can be linked to transparency: we need to get a handle on the implicit assumptions in machine learning.

- If the assumptions implicit in the machine learning methods are uncovered, we can relate them to earlier models.
- A clear insight into the assumptions will allow us to criticize the methods and explain the results.

We have to uncover the inductive assumptions in machine learning methods.

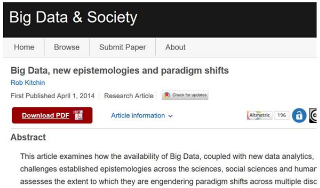
2 Concerns over reliability

Another motivation for making machine learning transparent derives from concerns over reliability.



CHRIS ANDERSON SCIENCE 06.23.08 12:00 PM

THE END OF THEORY: THE DATA DELUGE MAKES THE SCIENTIFIC METHOD OBSOLETE



Big Data, new epistemologies and paradigm shifts

Rob Kitchin

First Published April 1, 2014 | Research Article

Abstract

This article examines how the availability of Big Data, coupled with new data analytics, challenges established epistemologies across the sciences, social sciences and humar assesses the extent to which they are engendering paradigm shifts across multiple disc

Several machine learning researchers have proclaimed the "death of theory".

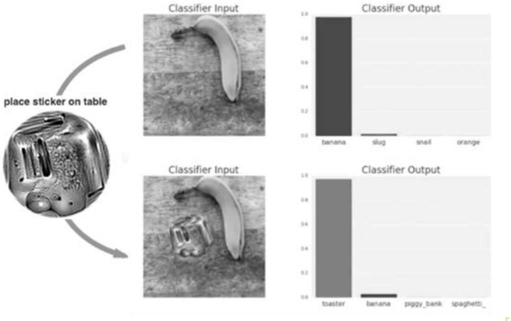
No free lunch

All inductive methods are in some way dependent on theoretical starting points: “there is no free lunch”.

If we have no control over the implicit assumptions of our methods, we do not know their conditions of applicability.

Inevitable inductive bias

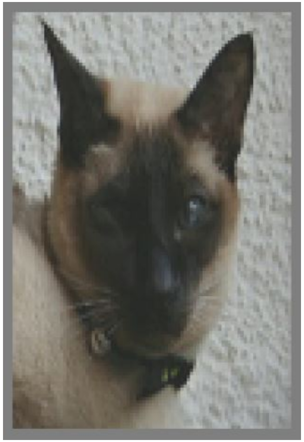
As illustrated by so-called adversarials, machine learning methods are vulnerable to highly unexpected error.



We have to gain insight into the inductive assumptions to gain control over the reasons for misfiring and “debug”.

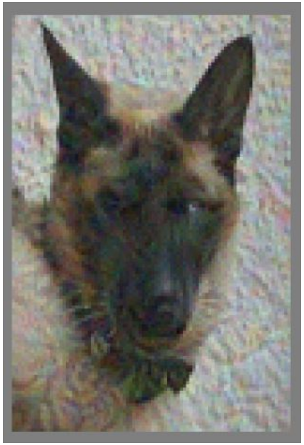
To illustrate adversarials. . .

What animal is this? Computer says “cat”.



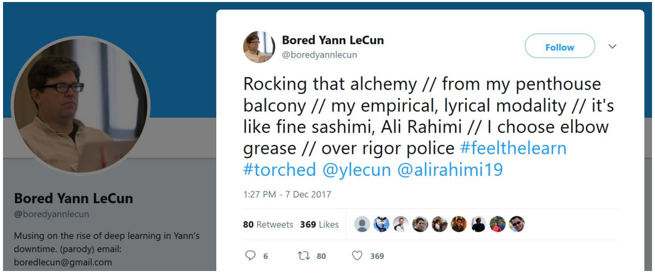
Adding a layer of noise

So what animal is this? Computer says “dog”.



At NIPS 2017

Rahimi sparked a fierce debate by deeming machine learning the “new alchemy” and calling for an active “rigor police”.



Bored Yann LeCun
@boredyannlecun

Musing on the rise of deep learning in Yann's downtime. (parody) email: boredlecun@gmail.com

Bored Yann LeCun
@boredyannlecun

Rocking that alchemy // from my penthouse balcony // my empirical, lyrical modality // it's like fine sashimi, Ali Rahimi // I choose elbow grease // over rigor police #feelthelearn #torched @ylecun @alirahimi19

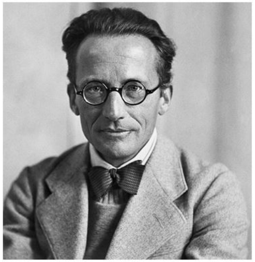
1:27 PM - 7 Dec 2017

80 Retweets 369 Likes

This debate shows parallels to a philosophical debate on the use of theoretical concepts.

“Anschaulichkeit”

The development of quantum mechanics offers an interesting example of the need for intelligibility.



Whether for epistemic, metaphysical or pragmatic reasons, scientists prefer theories that provide insights alongside predictions.

Wish list

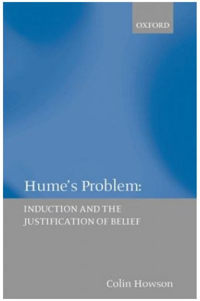
In sum, despite the attractiveness of theory-free machine learning methods, we want methods to. . .

- allow continuity in research,
- facilitate accountability,
- be understandable and communicable,
- have clear application criteria,
- avoid erratic mistakes.

For this we need clarity on the assumptions. How to reconstruct those?

Deploy Bayesian logic

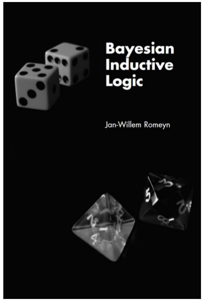
Following Howson’s view that probability theory is a logic, we can excavate the inductive assumptions that drive our inductive procedures by writing them down in a Bayesian format.


OXFORD

Hume’s Problem:

INDUCTION AND THE
JUSTIFICATION OF BELIEF

Colin Howson


Bayesian
Inductive
Logic

Jan-Willem Romeyn

In what follows I will offer several examples of this general idea.

3 First example: analogical predictions

Inductive logic is arguably a precursor of machine learning. Consider sampling pieces of fruit Q_i:

Carnapian predictions are made on the basis of data alone:

$$P(Q_{n+1} = \alpha | Q_1 \ldots Q_n) \;=\; \frac{n_\alpha + \lambda/k}{n + \lambda},$$

where the number of possible results $k = 4$ and we might choose $\lambda = k$.

Analogy effects

Carnap gradually admitted more flexibility in the prediction rules. A good example is analogical prediction, e.g.,

$$P(Q_{n+1} = a | Q_1 \ldots Q_n) \quad = \quad \frac{n_{\{a,c\}} + \mu/2}{n + \mu} \times \frac{n_a + \lambda/2}{n_{\{a,c\}} + \lambda}.$$

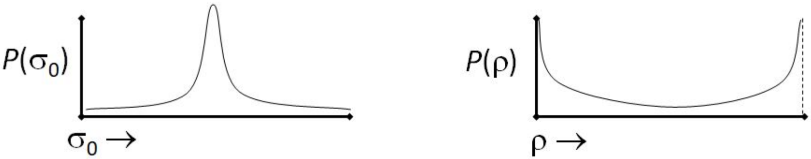
If $\mu < \lambda$, apples and bananas affect our expectation of cherries differently:

$$P(Q_{n+2} = c | Q_1 \ldots Q_n \wedge Q_{n+1} = a) \quad > \quad P(Q_{n+2} = c | Q_1 \ldots Q_n \wedge Q_{n+1} = b).$$

The literature offers numerous other systems that provide a handle on similarity in the data.

Using Bayesian statistics

Translating these prediction rules into Bayesian models is illuminating. We can redefine analogical prediction in fully Bayesian terms, by a prior over multinomial distributions: $P(H_\theta)$ where $\theta \in \langle \rho, \sigma_0, \sigma_1 \rangle$.



Here ρ is the probability for being round, and the σ's are the probabilities of having a stone conditional on being round or not.

Putnam’s curse

Notably, there is a striking parallel between adversarials and so-called unlearnable sequences in inductive logic.

- Putnam (1963) challenged Carnap’s project by constructing a sequence that, relative to a set of prediction rules, is not predictable.
- Once a rule assigns a high probability to an observation, the sequence will catch it by surprise and break the pattern.
- The formal learning theory developed after Putnam might shed light on the actively researched issue of adversarials in machine learning.

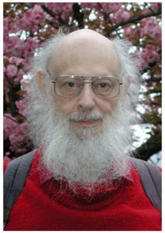
4 Uncovering inductive assumptions

Philosophy and statistics have seen many unsuccessful attempts to rid inductive inference from its theoretical starting points.

We can learn from these examples to inform our analysis of machine learning. Where did the implicit theoretical assumptions go to hide?

Universal prediction

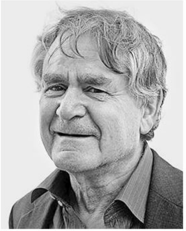
Sterkenburg (2017) offers an in-depth analysis of Solomonoff’s idea of universal prediction, i.e., of considering all possible data patterns in prediction.

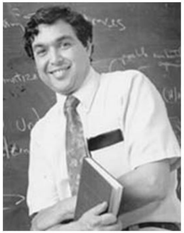


The predictions rest on the assumption of a highly skewed prior over all semi-computable measures. And in the end they fall prey to Putnam’s curse.

Fiducial argument

Fisher attempted to generate probabilistic conclusions about statistical hypotheses on the basis of data only.

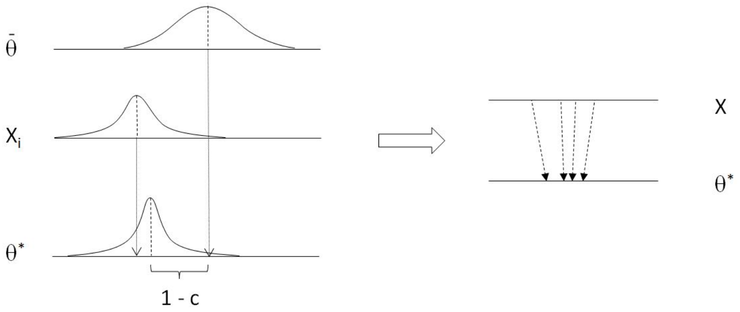




But. . . his argument rests on the assumption of an improper implicit prior, projected onto the hypotheses via a functional model.

Shrinkage estimators

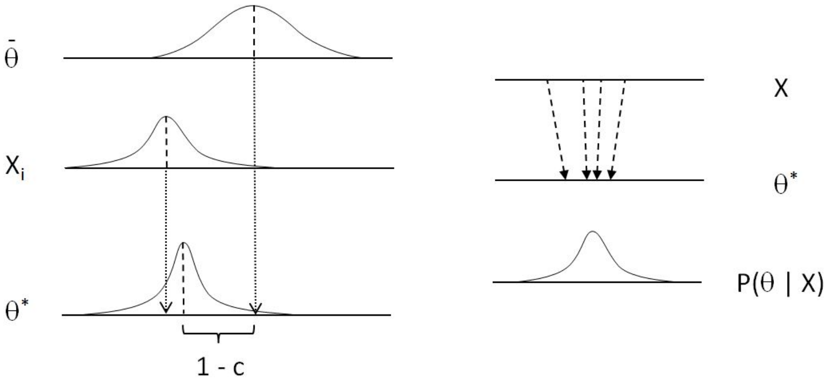
James and Stein (1957) derive that maximum likelihood estimators can be improved if we consider a collection of estimation problems.



As Efron and Morris (1977) show, the predictive improvement rests on an implicit empirical prior.

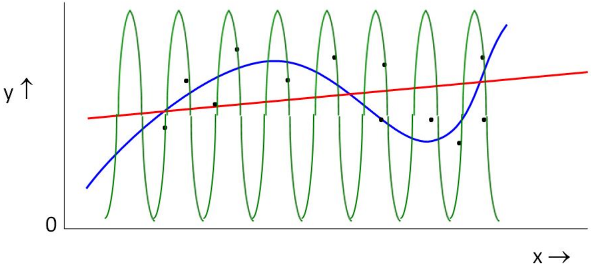
An empirical Bayesian model

Framed as a Bayesian method, Stein’s shrinkage factor approximates the Kalman filter. The nudge towards the grand mean is the result of the specific prior that we chose for θ.



Model complexity

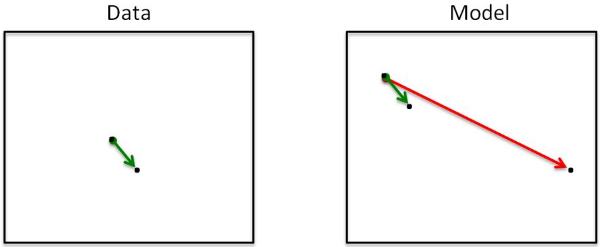
The sine model below offers a perfect fit while only using three free parameters.



This is an instant model selection hit! Fourier analysis trumps Taylor expansions.

Robustness and degeneracy

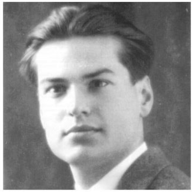
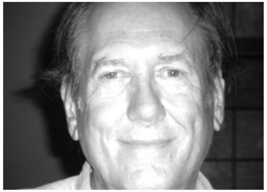
Curiously a small nudge in the data causes the best estimate to change radically. But the real problem is that the model is degenerate.



This will show up in a properly approximated marginal likelihood: the prior term will act as substantial penalty.

Bayesian logic

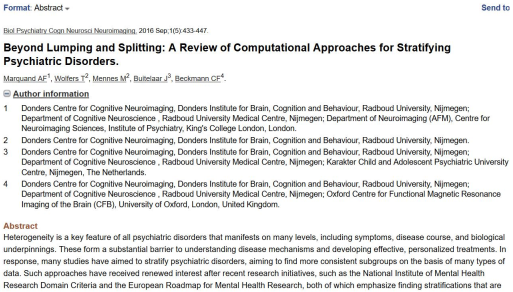
In all these cases the inductive assumptions are made explicit by converting the inductive procedures into Bayesian format.

The basis for this is in the view that the Bayesian format is a logic and as such epistemologically neutral.

5 Data-driven psychopathology

On to a scientific case study. . . Psychiatric classification and sub-typing can be assisted by automated clustering methods.



Biol Psychiatry Cogn Neurosci Neuroimaging. 2016 Sep;1(5):433-447.

Beyond Lumping and Splitting: A Review of Computational Approaches for Stratifying Psychiatric Disorders.

Marquand AF[1], Wolfers T[2], Mennes M[2], Buitelaar J[3], Beckmann CF[4].

Author information

1 Donders Centre for Cognitive Neuroimaging, Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen; Department of Cognitive Neuroscience , Radboud University Medical Centre, Nijmegen; Department of Neuroimaging (AFM), Centre for Neuroimaging Sciences, Institute of Psychiatry, King's College London, London.

2 Donders Centre for Cognitive Neuroimaging, Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen.

3 Donders Centre for Cognitive Neuroimaging, Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen; Department of Cognitive Neuroscience , Radboud University Medical Centre, Nijmegen; Karakter Child and Adolescent Psychiatric University Centre, Nijmegen, The Netherlands.

4 Donders Centre for Cognitive Neuroimaging, Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen; Department of Cognitive Neuroscience , Radboud University Medical Centre, Nijmegen; Oxford Centre for Functional Magnetic Resonance Imaging of the Brain (CFB), University of Oxford, London, United Kingdom.

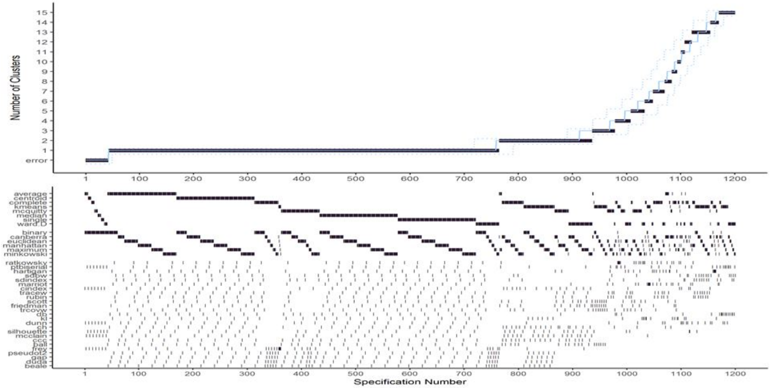
Abstract

Heterogeneity is a key feature of all psychiatric disorders that manifests on many levels, including symptoms, disease course, and biological underpinnings. These form a substantial barrier to understanding disease mechanisms and developing effective, personalized treatments. In response, many studies have aimed to stratify psychiatric disorders, aiming to find more consistent subgroups on the basis of many types of data. Such approaches have received renewed interest after recent research initiatives, such as the National Institute of Mental Health Research Domain Criteria and the European Roadmap for Mental Health Research, both of which emphasize finding stratifications that are

Do the methods identify patient groups that are distinct for the purpose of prediction and intervention?

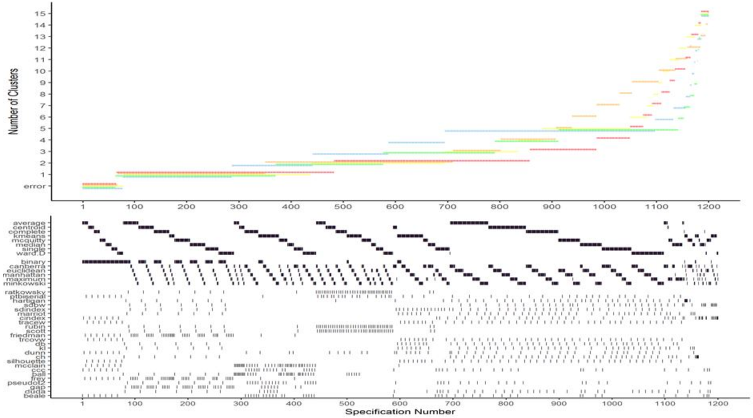
Specification curves

In a large comparison of clustering methods, Beijers et al. (manuscript) did not find much stability in the attempted clusterings.



Specification curves (continued)

When repeating the procedures for simulated data that were constructed to allow for easy detection, the same failures obtain.



The defects of automated clustering

We must not write off the use of data-driven methods in psychopathology but there are serious problems.

- There is wide variation and little overlap among the results of clustering subtypes of mental disorders.
- The comparison does not point to any particular specifications as being most adequate.
- The theoretical choices do not relate to the clustering outcomes determined by them in a conspicuous way.
- Variance, noise variables, and outliers all contribute to the failure of the clustering.

Towards a clarification of clustering

Again, framing the clustering methods in terms of a Bayesian logic helps us to see what assumptions might motivate the clustering method.

- Preliminary work suggests useful parallels between clustering and least-squares curve-fitting.
- Increasing the number of clusters is similar to increasing the number of parameters describing a family of curves.
- Any automated clustering method can be replicated by a hierarchical Bayesian model with distributional assumptions on the nature of a cluster.

6 The logic of data science discovery

Philosophy of science can help to introduce machine learning methods into science in a responsible way.

- Machine learning will very likely transform our sciences so we will have to focus attention there.
- Preliminary studies suggest that the outcomes of machine learning methods suffer from failures of robustness: unless assisted, they overfit.
- To improve on the assistance, our primary goal should be to identify the assumptions inherent in machine learning.

Making the assumptions explicit

The foregoing suggests how we can uncover inductive assumptions inherent in the new machine learning methods.

The idea is that we can identify modeling assumptions by translating the machine learning into Bayesian logic.

Why again?

Uncovering the assumptions of machine learning is an important task for the philosophy of science.

- It will help to integrate the new methods into existing and more theoretical approaches.
- Similarly it will improve on the communicability and public acceptance of machine learning results.
- And it will make it easier to hold researchers accountable and critically scrutinize their conclusions.
- Most importantly, it will help to apply methods correctly and guard against unreliable inferences.

Beyond logic

There are more assumptions to take into account though, and they are not all covered by logical analysis.

- Machine learning also relies on how the sample space and the space of theoretical possibilities is constructed.
- Many machine learning methods include a form of model selection, and thereby a decision procedure, over and above model evaluation.
- The application of machine learning methods involves interpretations of their results.

Thanks for your attention

Help from Talitha Anthonio and Lian Beijers is gratefully acknowledged.

Slides of the talk will be available at `http://www.philos.rug.nl/~romeyn`.

For comments and questions, email j.w.romeijn@rug.nl.