PhilML conference
Tübingen 2022

Machine Learning

or: the Return of Instrumentalism

★

Jan-Willem Romeijn
University of Groningen

Plan of talk

1. Data-driven science
2. Clustering in psychopathology
3. Instrumentalism
4. Insights from inductive logic
5. Application to automated clustering
6. Instrumentalism about data science

1 Data-driven science

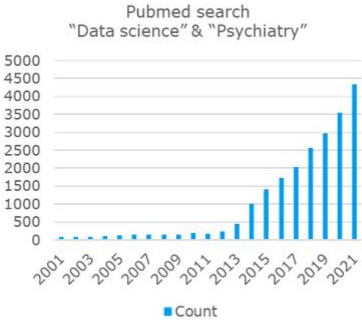
There are many examples of data-driven methods in the sciences, for prediction and automated model construction:

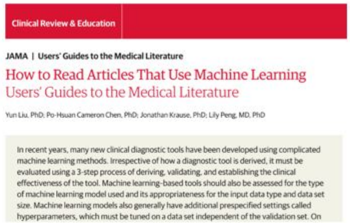
- Astronomers apply automated labelling of galaxies to generate new typologies and identify objects of interest.
- Biomedical researchers employ methods of automated causal discovery to identify mechanisms of gene expression in the cell.
- Psychiatrists use hierarchical clustering to identify subtypes of heterogeneous diseases like depression.

What does this methodological development do to these sciences?

Rapid uptake

Psychiatric science is seeing a rapid uptake of new data-scientific tools.




Clinical Review & Education

JAMA | Users' Guides to the Medical Literature

How to Read Articles That Use Machine Learning
Users' Guides to the Medical Literature

Yun Liu, PhD; Po-Hsuan Cameron Chen, PhD; Jonathan Krause, PhD; Lily Peng, MD, PhD

In recent years, many new clinical diagnostic tools have been developed using complicated machine learning methods. Irrespective of how a diagnostic tool is derived, it must be evaluated using a 3-step process of deriving, validating, and establishing the clinical effectiveness of the tool. Machine learning-based tools should also be assessed for the type of machine learning model used and its appropriateness for the input data type and data set size. Machine learning models also generally have additional prespecified settings called hyperparameters, which must be tuned on a data set independent of the validation set. On

This uptake goes hand-in-hand with increased interest in methodological guidelines for these methods.

Popular reception

The public perception of science is heavily impacted by new data science tools. Personalized evidence-based medicine seems very promising.

At the same time the concerns over the accountability and intelligibility of evidence-based decision making are growing.

2 Clustering in psychopathology

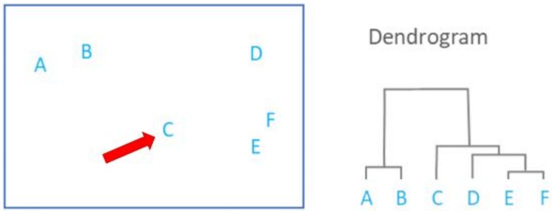
Psychiatric classification and sub-typing is assisted by automated clustering methods in the clinic.

Patient C.

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

How does it work?

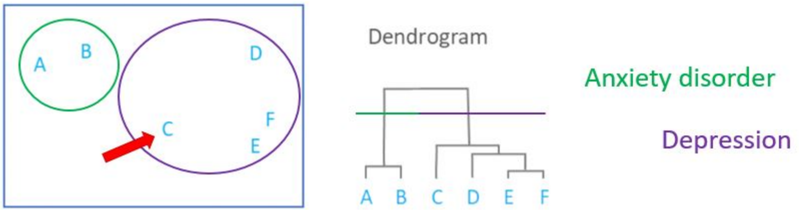
Here is a quick tour past the hierarchical clustering techniques that automated classification is based on.



The starting point is a space of patient characteristics and a tree structure expressing the proximity of the individuals in it.

Generating a classification

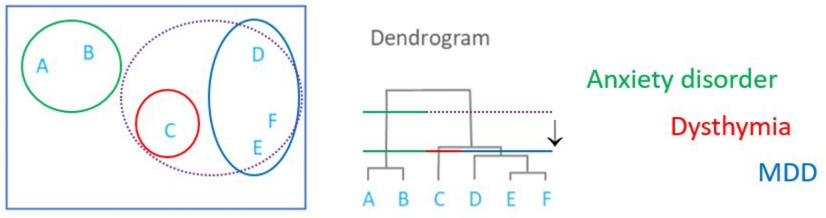
By choosing a certain granularity for the clusters we obtain a labelling for the patients.



This granularity can be determined by the classification system itelf, somewhat akin to model selection methods.

And generating a different one

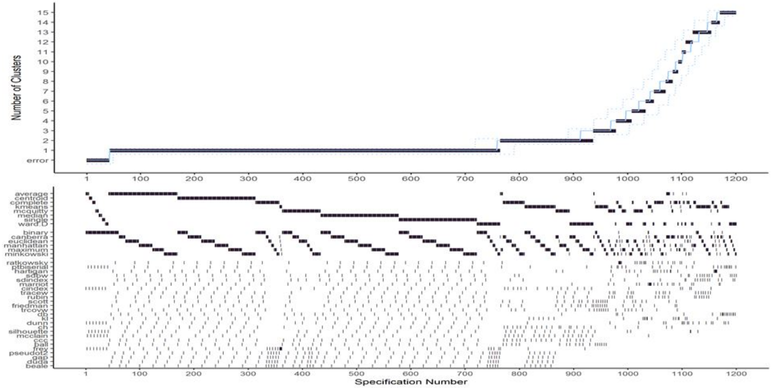
The resulting classification depends on many factors and parameter settings within the system.



Such settings are eventually, though often unreflectively, determined by the users of the system.

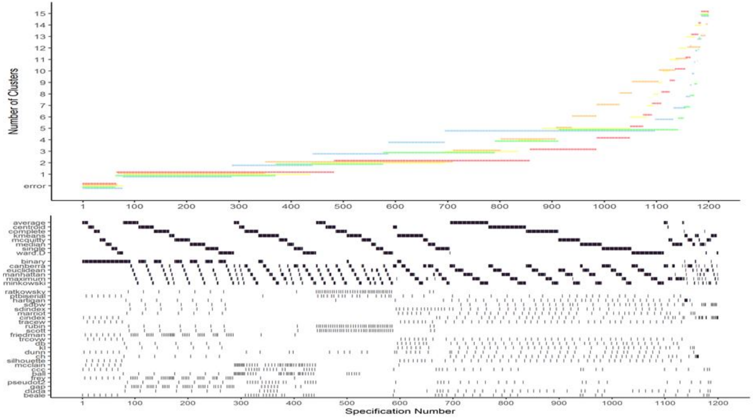
Specification curves

In a large comparison of clustering methods, Beijers et al. (2021) 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.



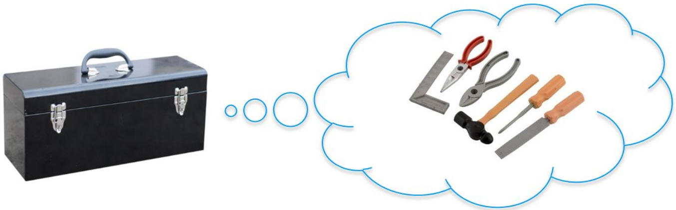
The unreliability of automated clustering

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 comparisons do not point to any particular settings 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.

3 Instrumentalism

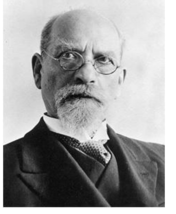
The “data science revolution” is arguably a rerun of a much older instrumentalist idea of theory-free science.



It expresses the logical empiricist viewpoint that we can build up scientific knowledge from empirical facts and their logical relations.

Empiricist social and medical science

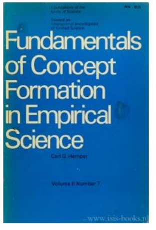
In search for a more solid scientific status, several “human” sciences disavowed theoretical development and turned to statistics.

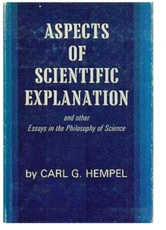


Popper’s critical rationalism reinforced this by obscuring the context of discovery and fostering a falsification mindset.

Reappraisal of theory

Logical empiricism developed towards giving a more prominent role to theoretical structures, modestly started in the work of Hempel.





We can hear the echo of these developments in the machine learning community, in particular in the call for explainable AI.

Theory crisis

Recently the absence of theoretical structure is hotly debated in psychological science, as something that obstructs the advance of their science.

- Epistemic freezing: once psychological constructs have been operationalized, we do not iterate to improve on their conceptualization.
- Testing myopia: psychological methodology focuses heavily on testing against data, thereby ignoring the hypotheses formation phase.
- Data fixation: the explanans is located at the level of data and not at the level of phenomena.

4 Insights from inductive logic

Carnapian inductive logic is arguably a precursor of machine learning: data are the only input. Consider sampling pieces of fruit Q_i:

Carnapian predictions are made on the basis of data alone:

$$P(Q_{n+1} = a | Q_1 \dots Q_n) = \frac{n_a + \lambda/k}{n + \lambda},$$

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

Gruesome predicates

In the received view, the inductive logic program was dealt a severe blow by Goodman's so-called new riddle of induction (1955).

The crux of the argument is that Carnap's prediction machinery can only be run after we have chosen our projectible predicates.

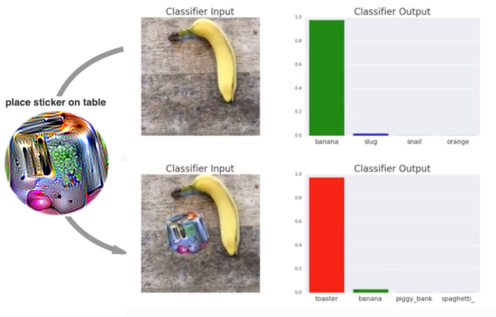
Unlearnable sequences

The more definitive blow, in my view, came from Putnam (1963), who devised an argument against the possibility of a universal learning machine.

- An adversarial data sequence is a sequence that is constructed to make the prediction system fail.
- Namely, anytime the system latches onto a pattern and assigns high probability to its continuation, the pattern in the sequence is broken.
- All predictive systems are restricted in their sensitivity to patterns, and therefore all are vulnerable to adversarial data.

Adversarials in machine learning

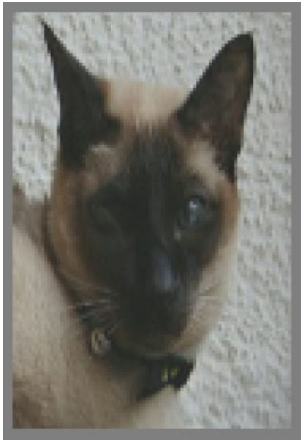
As discussed earlier today, machine learning methods are also vulnerable to adversarial data.



The lessons for current-day machine learning are the same as those for Carnap's program: pay attention to inductive bias.

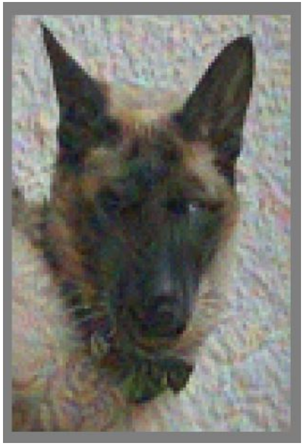
To illustrate adversarials. . .

What animal is this? Your brain says “cat”.



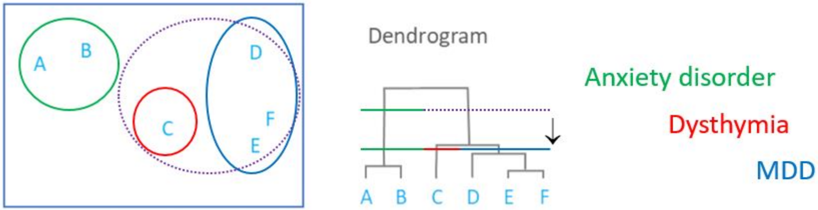
Adding a layer of noise

So what animal is this? Your brain says “dog”.



5 Application to automated clustering

Insights on the problems of instrumentalism can be readily applied to machine learning, in particular to automated clustering methods.



They arguably reveal why the automated psychiatric classification systems fail to deliver: their assumptions are not accounted for.

The lesson from Goodman

Many inductive assumptions enter into the predictive system through data construction. An instrumentalist cannot properly account for this.

Is our data ever rich enough to capture clinical reality? Psychiatric science needs of Ryle’s and Geertz’s notion of thick description.

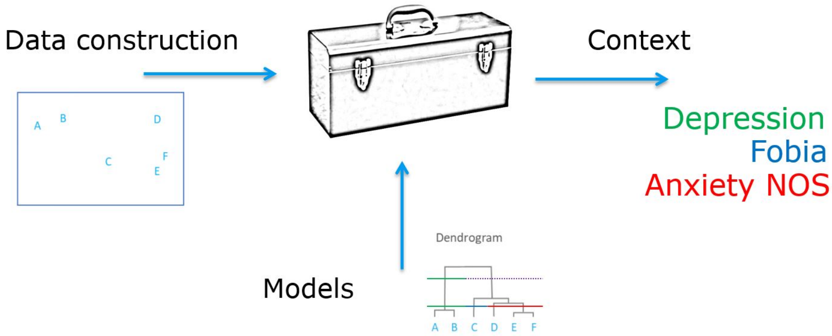
The lesson from Putnam

Framing the clustering methods in terms of a statistical inference helps to see what assumptions and biases drive the clustering.

- There are useful parallels between clustering and curve-fitting. Statistical model selection is central to both.
- Specifically, increasing the number of clusters is similar to increasing the number of parameters describing a family of curves.
- Automated clustering method can be replicated by a hierarchical Bayesian model with distributional assumptions on the nature of a cluster.

Making the black box transparent

To fully grasp the inductive assumptions that drive the automated clustering, we need to get the whole application process into view.



Besides data construction and modelling, we also need to clarify how the results are interpreted and used.

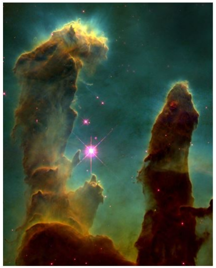
6 Instrumentalism about data science

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

- Data science will very likely transform our sciences so we will have to focus attention there.
- Preliminary studies suggest that the outcomes of these methods suffer from failures of reliability.
- To improve on the assistance, our primary goal should be to make the data science methods transparent.

The double curse of instrumentalism

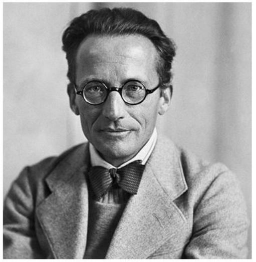
An instrumentalist attitude towards data science methods goes against the goal of transparency and thus hampers responsible data science use.



It is all the more dangerous for a science like psychiatry because it is low on theory and it has high impact on people’s lives.

“Anschaulichkeit”

The development of quantum mechanics offers an interesting parallel to this need for intelligibility.



Whether for epistemic, pragmatic, or even metaphysical reasons, theories that provide understanding alongside predictions are preferable.

Uncovering inductive assumptions

Statistical science has seen many unsuccessful attempts to rid inductive inference from its theoretical starting points.

We can learn from these attempts, and from the philosophy of statistics about them, to inform our analysis of machine learning.

Translation to statistical inference

De Finetti style representation theorems suggests how to uncover inductive assumptions inherent in machine learning methods.

- Identify modeling assumptions by translating between predictive systems and Bayesian statistical inference.
- Consider the assumptions inherent in how the sample space and the space of hypotheses is constructed or developed.
- Connect such assumptions explicitly to what goes into the Bayesian model: a family of distributions and a prior probability.

Thanks for your attention

Help from Lian Beijers and Hanna van Loo 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.