Schizophrenia Bulletin vol. 44 no. 5 pp. 939–941, 2018
doi:10.1093/schbul/sby101
Advance Access publication July 9, 2018

EDITORIAL

Will Machine Learning Enable Us to Finally Cut the Gordian Knot of Schizophrenia

Neeraj Tandon[1] and Rajiv Tandon[*,1]

[1]Department of Psychiatry, University of Florida College of Medicine, Gainesville, FL

*To whom correspondence should be addressed; tel: 352-548-6454, fax: 352-548-1627, e-mail: tandon@ufl.edu

The field of schizophrenia research is at a crossroads. On the one hand, relatively little progress has been made in elucidating its fundamental nature or in developing more effective treatments, leading to increasing calls for the death of this construct[1] and its immediate replacement by one of several proposed alternatives. On the other hand, there is no consensus about which of these alternatives should replace it as none of them have been found to better explain the set of facts associated with schizophrenia.[2] Can we make transformational advances in our comprehension of human brain function and apply that understanding into a more accurate concept of schizophrenia? Multibillion-dollar research initiatives such as the US-based Brain Research through Advancing Innovative Neurotechnologies and the Europe-based Human Brain Project hope to revolutionize our appreciation of how the human brain works. A prerequisite to these efforts is the ability to integrate and analyze "big data," enabled by the exponential increase in the capacity of computer systems to store and process data. This enhanced capability has spawned great excitement in the overlapping fields of computational psychiatry and network neuroscience,[3] exemplified by the steep growth of scientific publications in the area. This trend is also reflected in the field of schizophrenia research;[4] for example, the number of publications on machine learning (ML) in this Journal has increased from an average of 1/year between 2004 and 2013 to 9 in 2017 and 15 this year. The 5 articles on the topic in this issue of the Journal illustrate both the promise and challenges in the application of ML methods to the study of schizophrenia.[5–9]

Arthur Samuel, who coined the term machine learning, defined it as "a field of study that gives computers the ability to learn without being explicitly programmed."[10] The machine (computer) reveals relationships between different variables and categorizes individuals based on these relations and/or uncovers pathways between different variables without being specifically instructed—the organization of the data is "naturally" derived directly from the input data rather than being limited or "distorted" by some preconceived idea of the relationships between the data elements. In addition, in contrast to traditional statistical methods, the current capacity of the machines also allows dissection of multilayered complex relationships among the variables. These capabilities of ML are evident in the studies described in the 5 articles on the topic in this issue. (Input variables are analogous to independent variables and the output variable is analogous to the dependent variable).

Does ML then reveal the true nature of relationships, unconstrained by any bias or human influence? The answer is an unequivocal No. As illustrated in figure 1, there are several points of potential "distortion" in the process of ML and its application:

1. At the point of data collection with the choice of the sample on which to collect information, what variables to collect information on, how to collect the information, etc.
2. Raw data are generally not input as collected. Instead, they are processed or "cleaned up" in some way and the manner of this processing can "distort" the output.
3. Some specifications are often placed on the model or algorithm and each ML method introduces some constraints.
4. Furthermore, the interpretation of the function derived from ML may or may not be useful or relevant. It may not be comprehensible.
5. The model or algorithm generated may or may not be replicable or reproducible. Each model comes with a cost-variance balance wherein there is a trade-off between precision and generalizability.

Each of these constraints, demonstrated in the studies described in the 5 articles on ML in this issue, pervade all ML studies in schizophrenia. They are not, however, unique to ML studies and are evident in non-ML studies

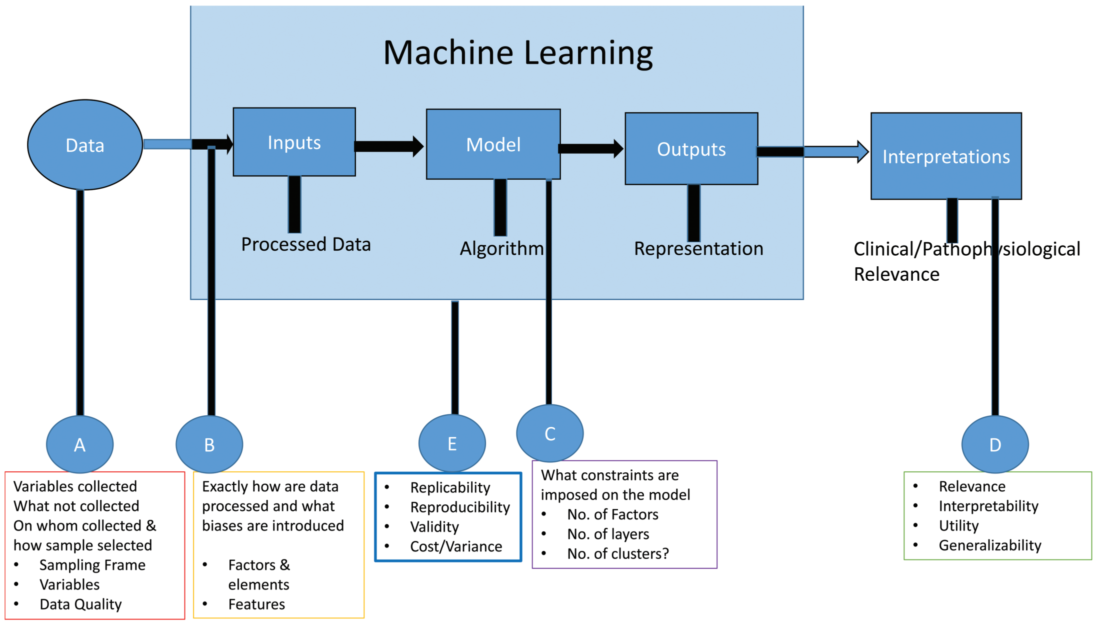

Fig. 1. The nature of machine learning.

as well. What needs to be recognized is that while techniques of ML are powerful tools, they cannot compensate for constraints imposed in various stages in their application (figure 1). A better appreciation of the nature of ML and its more rigorous application will better position our field in harnessing the enormous potential of ML in elucidating the true nature of schizophrenia. ML researchers should provide greater detail about their specific methods at each stage (figure 1) and discuss the impact of their chosen parameters on the interpretation of the resultant model. In addition, because of the black box nature of model generation in ML, implicit assumptions underlying its specific application in a study should be spelled out. Finally, the specific question being addressed in any ML study should be clearly defined and the model that is generated should align with this purpose and its utility/relevance should be continually evaluated. Readers need to develop a better understanding of the nature and types of ML. They should recognize that ML reveals patterns of association and therefore always generates models or hypotheses that need to be specifically tested. Additionally, ensuring greater literacy about ML in the field at large is a critical need.

Perhaps the computational tools of ML will enable a finer dissection of the neurobiological underpinnings of schizophrenia and thereby help break the current impasse.

In schizophrenia research, as exemplified in this Journal edition, researchers have begun to seize on the opportunities presented by ML to advance knowledge and decision-making. Although our field's exuberance about the promise of ML is justified, discipline is warranted. The power of ML is such that its application to varied schizophrenia datasets will certainly generate a plethora of findings. Whether these findings contribute to a more valid construct of schizophrenia or merely add to the many unused bricks already strewn in the schizophrenia brickyard[11,12] significantly depends on how we use ML in our research efforts. In all the hurry and hubbub, may we never lose sight of the prize: our goal is a deeper understanding of the pathophysiology of schizophrenia leading to improved treatment.

References

1. Goluksuz S, van Os J. The slow death of the concept of schizophrenia and the painful birth of the psychosis spectrum. *Psychol Med*. 2018;48:229–244.
2. Tandon R, Keshavan MS, Nasrallah HA. Schizophrenia, "Just the Facts": what we know in 2008 part 1: overview. *Schizophr Res*. 2008;100:4–19.
3. Redish AD, Gordon JA. *Computational Psychiatry: New Perspectives on Mental Illness*. Boston: MIT Press; 2016.
4. Krystal JH, Murray JD, Chekroud AM, et al. Computational psychiatry and the challenge of schizophrenia. *Schizophr Bull*. 2017;43:473–475.
5. Cui LB, Liu L, Wang HN, et al. Disease definition for schizophrenia by functional connectivity using radiomics strategy [published online ahead of print February 17, 2018]. *Schizophr Bull*. doi: 10.1093/schbul/sby007.
6. Zheutlin AB, Chekroud AM, Polimanti R, et al. Multivariate pattern analysis of genotype-phenotype relationships in schizophrenia [published online ahead of print March 9, 2018]. *Schizophr Bull*. doi: 10.1093/schbul/sby005.
7. Dwyer DB, Cabral C, Kambeitz-Ilankovic L, et al. Brain subtyping enhances the neuroanatomical discrimination of schizophrenia [published online ahead of print February 26, 2018]. *Schizophr Bull*. doi: 10.1093/schbul/sby008.
8. Rozycki M, Satterthwaite TD, Koutsouleris N, et al. Multisite machine learning analysis provides a robust structural imaging signature of schizophrenia detectable across diverse patient

populations and within individuals [published online ahead of print November 24, 2017]. *Schizophr Bull*. doi: 10.1093/schbul/sbx137.

9. Koutsouleris N, Wobrock T, Guse B, et al. Predicting response of repetitive transcranial magnetic stimulation in patients with schizophrenia using structural magnetic resonance imaging: a multisite study. *Schizophr Bull*. 2018;This issue.
10. Samuel AL. Some studies in machine learning using the game of checkers. *IBM J Res Dev*. 1959;3:535–554.
11. Forscher BK. Chaos in the brickyard. *Science*. 1963;142:339.
12. Tandon R. Moving beyond findings: concepts and model-building in schizophrenia. *J Psychiatr Res*. 1999;33:467–471.