Rethinking the Role of Machine Learning in Biomedical Science

Konrad Paul Kording, Ari Benjamin, Roozbeh Farhoodi, Joshua Glaser

While the direct goal of biological modeling is to describe data, modeling ultimately aims to find ways of fixing systems and understanding systems' objectives, algorithms, and mechanisms. Thanks to engineering applications, machine learning is a booming field that allows modeling data extremely well, typically without using explicit assumptions about the modeled system. Machine learning can typically better describe the data than biomedical models and thus provides both engineering solutions and an essential benchmark. It can also be a tool to aid understanding. Using examples from neuroscience, we highlight the contributions, both realized and potential, of machine learning. Machine learning is becoming easy to use and should be a critical tool across the full spectrum of biomedical questions.

Introduction

The goal of nearly all of computational biology is to numerically describe a system, which is often quantified as the explained variance. In some cases, we only care about the explained variance, e.g. if we need to make predictions. But in most cases, just describing the data successfully is not sufficient and there has been much discussion of such objectives in the neuroscience community[1, 2]. In some cases we want the model to inform how we can fix things; in other words we want it to predict what would happen if we did interventions. In some cases we want to understand the data as optimizing an objective. For example, we might ask if the intricate folds of the brain minimize wiring length[3]. In some cases we want to understand the system as an algorithm, e.g. which algorithms the brain uses to learn[4]. Probably most commonly, we want to understand underlying mechanisms. For example, we may want to know how action potentials are enabled by interactions between voltage dependent ion channels[5]. The typical model cares not only about numerically describing data, but also about other objectives of the researcher.

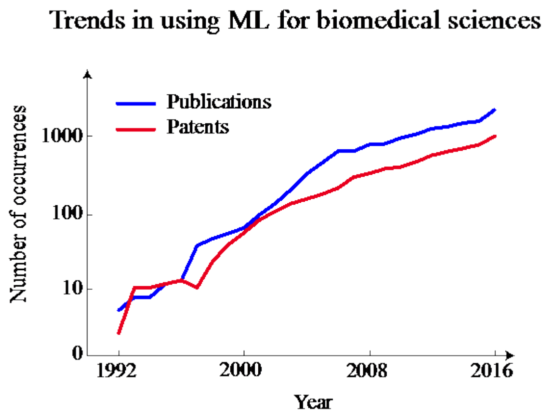

Figure 1: The trend towards using machine learning for biomedical sciences: The use of machine learning (ML) in biomedical sciences is increasing. Publication data (blue) was collected from the Semantic Scholar website using the keywords "biomedical" and "machine learning." The patent data (red) was collected from Google Patents using the same keywords.

So far, progress in the modeling field comes mostly from human insights into systems. People think about the involved components, conceptualize the system's behavior, and then build a model based on their intuitive insights. This is being done for neurons[2], molecules[6], and the immune system[7]. Across biomedical research scientists are starting to use computational models to both describe data, and to specify the underlying principles. However, understanding such complex systems is extremely difficult, after all our intuition is bound to be incomplete in systems with many nonlinearly interacting pieces[8].

Machine learning, or more precisely supervised learning, is a radically different way of approaching modeling, and it relies on minimal human insight. At its heart, supervised machine learning assumes that the relationship between the measured variables and those to be predicted is in some sense simple[9]; it assumes things like smoothness, sparseness, or invariances. Typical supervised algorithms get vectors of features as inputs and produce predictions as outputs. Machine learning techniques then mostly differ by the nature of the simplicity of the function they use for predicting[10]. As opposed to assuming an explicit model about the relationship of variables, machine learning techniques rather assume simplicity.

The field of machine learning is undergoing a revolution and has moved from a niche discipline to a major driver of economic activity over the last couple of decades. Indeed, the number of papers and patents in biomedical research has grown exponentially over the last decade (Figure 1). Progress in machine learning is revolutionizing speech-to-text, web-search, and countless other areas of economical importance. The influx of talent into this field has led to massive improvements of algorithm performance, allowing computers to outperform humans at tasks such as image recognition[11] and playing go[12]. These developments in machine learning promise to make it be an important tool in biomedical research.

The time has come for machine learning to impact biomedical research

There are many kinds of questions that can be answered well using machine learning techniques. In some cases we care about the predictions themselves, e.g. we may ask if a drug will cure a given cancer. In other cases, it sets a benchmark, e.g. how much worse is our human-thought-out model relative to what may be possible. In yet other cases machine learning may help us understand a system, e.g. by telling us if information is contained in a response or which variables are shared between components of a system. The breadth of potential uses of machine learning for biomedical research has not yet been realized.

The standard use case for machine learning is prediction. A typical problem in medicine may be the detection of mood. We have measured data about everyday behaviors, e.g. when patients wake up or how much they exercise. We want to estimate how their mood is changing. Many algorithms have been used to solve such problems[13]. A typical problem in neuroscience is neural decoding[14]. We have measurements from the brain of a subject that cannot move its body. We want to estimate which movements the subject wants to do, so that we can have a robot execute this movement. Many algorithms have been developed to solve such problems[15-17]. General purpose machine learning tends to do extremely well[18]. Computationally similar problems exist all over biomedical research including cancer[19], preventive medicine[20], and medical diagnostics[21]. In such problems, we do not care beyond the quality of the predictions. After all, that is what the application asks for. Many applications of machine learning in engineering only care about the typical error size.

In many other cases, however, we do not just want to describe and predict data but we want to be able to produce models that we can understand and teach to our students. In that case, machine learning can still be extremely useful by providing a benchmark. A general problem when evaluating models is that it is hard to know how much of the model's errors are due to noise and how much is due to the model being insufficient. Machine learning is generally good at making predictions, and could thus provide close to an upper bound for the models we produce ourselves. If a human-generated model does much worse than the machine learning benchmark, we might expect that we are missing important principles or that our modeling is misguided. If, on the other hand, a model based on human intuition is very close to the machine learning precision, we should be more convinced that the posited concepts are, indeed, meaningful. For example, when we predict mood on mobile phone data, we may build our model on psychological insights. Or when we describe how measured factors affect neural activities, we may use a linear-nonlinear-poisson model [22]. However, how do we know if our model is meaningful or is missing important aspects? Machine learning benchmarking can answer those questions[23]. Benchmarks are necessary for the interpretation of models.

In yet other cases, machine learning can directly help understanding. One important question that is often being asked is if a system carries information about some variables. But we may not really know about the relation; it may be linear or nonlinear. In such cases machine learning allows asking if information is contained in a signal without us having to specify the exact nature of the relationship. Another important question is about the information shared between two parts of a system. For example we may ask which aspects of the world (high dimensional) is shared with which aspects of the brain (also high dimensional). Machine learning allows asking such questions in a well defined way[24, 25]. For many questions in biology, machine learning promises to allow new approaches to generate biological understanding.

Machine learning is becoming a necessity for ever-growing datasets

Datasets are rapidly growing and becoming more complex. In neuroscience, the number of simultaneously recorded neurons is growing exponentially[26]. Similarly, the amount of electronic health record data is increasing rapidly. In fact, the overall biomedical literature is equally growing extremely fast. At the same time, the complexity of datasets is rapidly increasing with modern datasets often being multimodal and multifaceted[27]. There are several ways in which these changes in datasets will produce new problems in modeling.

First, humans are not very good at thinking about complex datasets. We can only consider a small hypothesis space. But in biology, as opposed to physics, we have good reasons to assume that truly meaningful models will be rather complex[28]. While humans will correctly see some structure in the data, they will simultaneously miss much of the real existing structure. It could be argued that it is nearly impossible for humans to intuit models of complex biological systems.

Second, nonlinearity and recurrence makes modeling complex systems much harder[28]. In this regime, it is hard to both explain complexity and ensure your model will fail if the causal structure is dissimilar to your model. In the limit of full-cell interactions or full-brain modeling, it seems implausible that we could design models that are poised at this delicate limit.

Finally, in the case of the large complex systems that are typical in biology, a major problem is that we do not understand the space of meaningful models. We can construct many models that will explain some portion of the variance, but it does not mean these models actually explain the mechanism. Comparing models can thus be pointless, when the models being considered are outside the space of meaningful models. In other words, we are effectively comparing models we do not believe with other models that we do not believe. We may have doubts that science can advance meaningfully in this manner.

Given all these arguments, we may believe that the typical non-machine-learning based approach to science that we learned from the physicists can only partially succeed. Any reasonably small number of principles can only describe some part of the overall variance, potentially a relatively small part. It is unclear how far the typical approach taken in biomedical research that draws on concepts of necessity and sufficiency can succeed at allowing us to understand the bulk of activity in complex interacting systems[29]. Machine learning covers a very different part of the space of relevant models. It has the potential to describe a very large part of the variance. At the same time, it has the drawback of being hard to communicate. Importantly, however, it covers a rather complementary part of the space of models and, as such, promises to become a central tool.

The use of ML increasingly does not require specialist knowledge

There are countless approaches in machine learning, certainly more than the typical biomedical scientist wants to know. Kernel based systems such as support vector machines are built on the idea of regulating model complexity[10]. Neural networks are built on the idea of hierarchical representations[30]. Random forests are built on the idea of having many weak learners[31]. One could easily fill books with all the knowledge that we have about machine learning techniques. However, using these techniques has become very simple. Training a machine learning technique basically just requires a matrix of training features and a vector of the known labels. As such, given the availability of the right software packages[32] we generally only need a few lines of code to train any machine learning system.

Moreover, techniques called ensemble methods obviate the need to even choose a single machine learning technique [33]. The idea is that a system can run all machine learning techniques and then combine the predictions of these using yet another machine learning technique. Such approaches generally win machine learning competitions (e.g. kaggle.com).

Furthermore, there is a new trend in machine learning that has developed rapidly over the last few years called automatic machine learning[34]. The idea is that most machine learning experts will do similar things when they approach machine learning. They will choose one of a number of methods (or all of them if they use ensembling). Then they will optimize the hyperparameters of those techniques. They may also optimize the feature representation. Long story short, while this can take a person a significant amount of time via trial and error, the pipeline is relatively standard and can be automated. And several new packages allow automating some or all of the process (e.g. https://github.com/automl/auto-sklearn, http://www.cs.ubc.ca/labs/beta/Projects/autoweka/, https://github.com/KordingLab/spykesML, https://github.com/KordingLab/Neural_Decoding). These developments are likely to pick up in speed over the next year or two making it less necessary for biomedical scientists to even know the details of the individual methods, freeing them to instead ask the real scientific questions that we want machine learning to answer.

Examples of state of the art machine learning in neuroscience

We want to go through two uses of machine learning approaches, predictions and benchmarking, using examples from neuroscience.

A first example is neural decoding. In neural decoding we want to estimate what a subject wants to do based on brain activity, for example to allow them to move an exoskeleton with their thoughts. A standard approach in the field is still the use of simple linear techniques, e.g. the Wiener Filter. In the Wiener filter, all signals from the past up to a time horizon are linearly combined to predict the output. There has been recently a lot of interest in improving such approaches using modern machine learning.

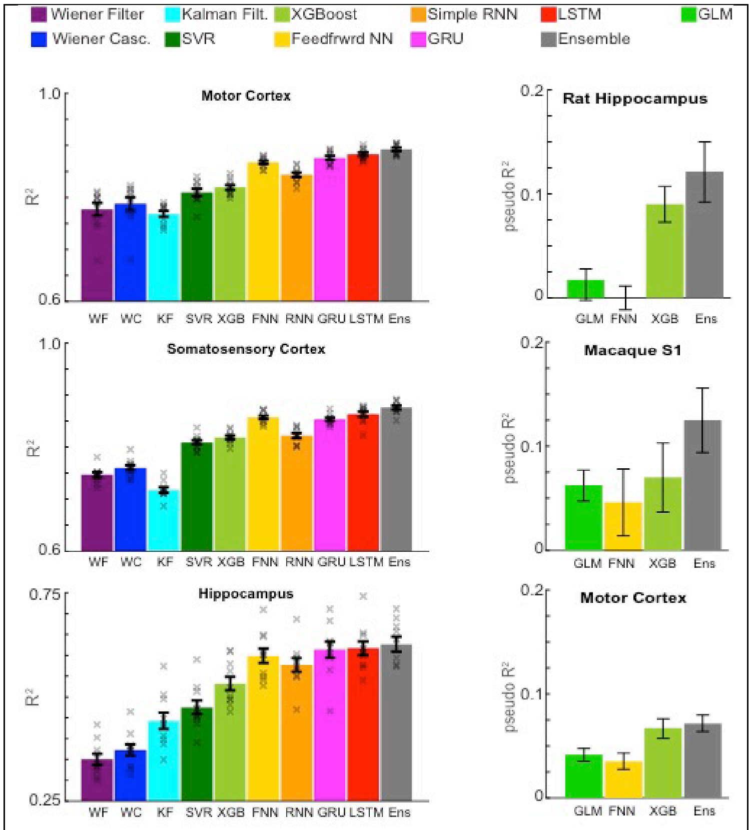

Figure 2: State of the art machine learning for neural decoding and encoding. Left, decoding, predicting state using signals from three brain areas using various machine learning techniques and the ensemble over them. Right, encoding, predicting spikes using signals from three brain areas. Data replotted with permission from [18, 23]

For many applications we just

want to have good performance. To analyze how much extra mileage we get out of using standard machine learning, we implemented many different approaches. This included the linear Wiener Filter, the nonlinear extension called Wiener Cascade, the Kalman filter, nonlinear support vector machines, extreme gradient boosted trees, and various flavors of neural networks (Figure 2). Ultimately, the modern neural network based techniques, as should be expected, did very well[18]. The combination of all of the techniques using ensemble methods performed even better. The same general phenomenon was seen when decoding from a number of different brain regions. Thus, when solving engineering problems, using standard machine learning should be the starting point.

In this sense, machine learning also sets a benchmark for other decoding approaches. When neuroscientists write decoding algorithms they are often inspired by their insights into the way the brain works[35]. However, without a comparison to modern machine learning, we cannot know if these insights work at all. As machine learning becomes automatic and easy to use, we argue that it should always be used as a benchmark for engineering applications.

A second example is neural encoding, or tuning curve analysis. In this case, we have signals from a neuron or a brain region and we want to know how this signal relates to variables in the outside world. Neuroscientists often believe that such a characterization allows insights into the role of a neuron in computation[8]. Typically, the neuroscientist chooses a model (often implicitly) based on prior knowledge about the brain. Subsequently the average signal is plotted as a function of the variables in the outside world. This approach implicitly assumes a linear model. But could we do better using machine learning?

For such applications it is impossible to know how good the tuning curve (encoding model) is at actually describing neural activity. In principle, input variables may affect the neuron’s activity in highly nonlinear ways. Machine learning allows us to test this hypothesis. Indeed, when we compared the generalized linear model (GLM)[22], which is mostly linear, it performed considerably worse than neural networks or extreme gradient boosted trees (Figure 2). And again, the combination of all the methods using ensemble techniques yielded the best results. Interestingly, despite the fact that the space was rather low dimensional, GLMs performed poorly relative to modern machine learning. This may suggest that the tuning curves measured by neuroscientists are rather poor at describing neurons in real world settings.

In this context, machine learning can conceptually contribute several aspects. First, it can detect that a variable is represented, even if there is no linear correlation. Second, it can set a benchmark that humans can strive for. Third, it offers a possibility of replacing the cartoon model that people have about neural computation with an, admittedly hard to interpret, complex alternative.

Discussion

The fact that machine learning is entering our toolbox has profound implications for all aspects of neuroscience and biomedical science in general.

Machine learning also changes the objectives of data collection. In traditional approaches, measuring many variables is unattractive as, through multiple comparison testing corrections, we will be unable to say much about each of them. In machine learning the situation is different. Many variables improve predictions, even if we cannot know which variables contribute to this. But this is not just a vacuous statement about information processing. It really reflects the fact that the brain and other biological systems are not simple, with few interactions, but highly recurrent and nonlinear. The assumption of simplicity used in biology is usually only a fanciful, if highly convenient, illusion.

Based on their intuitions, scientists are starting to fit rather complex models to biological data. Indeed, those complex models usually fit the data better than simpler models. However, a complex model based on a wrong idea may be able to fit the data extremely well. This issue is in some way negating the real advantage of interpretable models. A good fit does not mean that the model is right. For example, Lamarckian evolution explains a lot of data about species, and yet, it was based on a fundamentally misleading concept of causal transmission of traits. This problem affects human intuition based models. It is no problem for machine learning models that usually, by design, do not pretend to get at a meaningful causal interpretation.

In the space of biomedical modeling, traditional modeling and machine learning cover opposite corners. Traditional modeling leads to models that can be compactly communicated and taught, while explaining only a limited amount of variance. Machine learning modeling explains a lot of variance, but is in its typical forms exceptionally hard to communicate. Those two frameworks can inform one another and both should be used to their maximal possibility.

References

1. Marr D. *Vision: A computational approach*. San Francisco: Freeman & Co; 1982.
2. Dayan P, Abbott LF. *Theoretical neuroscience: Computational and mathematical modeling of neural systems*. MIT Press; 2001.
3. Van Essen DC. A tension-based theory of morphogenesis and compact wiring in the central nervous system. *Nature*. 1997;385:313
4. Marblestone AH, Greg Wayne, Konrad P. Kording. Toward an integration of deep learning and neuroscience. *Front Comput Neurosci*. 2016;10
5. Hodgkin ALaH, A.F. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of physiology*. 1952;117:500-544
6. Leszczynski J. Computational molecular biology. 1999;8
7. Petrovsky N, and Vladimir Brusic. Computational immunology: The coming of age. *Immunol. Cell Biol.*;80:248
8. Jonas E, Kording KP. Could a neuroscientist understand a microprocessor? *PLoS Comput. Biol.* 2017;13:e1005268

9. Wolpert DH. What the no free lunch theorems really mean; how to improve search algorithms. 2012
10. smola A, Schoelkopf B. *Learning with kernels*. MIT Press; 2002.
11. He K, Zhang, X., Ren, S., & Sun, J. . Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. *Proceedings of the IEEE international conference on computer vision*. 2015:1026-1034
12. Silver D, Huang, A., Maddison, C.J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M. and Dieleman, S. Mastering the game of go with deep neural networks and tree search. *Nature*. 2016;529:484-489
13. Wang R, Fanglin Chen, Zhenyu Chen, Tianxing Li, Gabriella Harari, Stefanie Tignor, Xia Zhou, Dror Ben-Zeev, and Andrew T. Campbell. Studentlife: Assessing mental health, academic performance and behavioral trends of college students using smartphones. *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 2014:3-14
14. Velliste M, Perel, S., Spalding, M. C., Whitford, A. S., Schwartz, A. B. Cortical control of a prosthetic arm for self-feeding. *Nature*. 2008;453:1098
15. Yu BM, Kemere C, Santhanam G, Afshar A, Ryu SI, Meng TH, et al. Mixture of trajectory models for neural decoding of goal-directed movements. *J. Neurophysiol.* 2007;97:3763-3780
16. Truccolo W, Eden UT, Fellows MR, Donoghue JP, Brown EN. A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects. *J. Neurophysiol.* 2005;93:1074-1089
17. Corbett EA, Perreault EJ, Kording KP. Decoding with limited neural data: A mixture of time-warped trajectory models for directional reaches. *J. Neural Eng.* 2012;9:036002
18. Glaser JI, Chowdhury, R. H., Perich, M. G., Miller, L. E., Kording, K. P. . Machine learning for neural decoding. *arXiv preprint arXiv*. 2017:arXiv:1708.00909
19. Kourou K, Exarchos, T. P., Exarchos, K. P., Karamouzis, M. V., Fotiadis, D. I. Machine learning applications in cancer prognosis and prediction. *Computational and structural biotechnology journal*. 2015;13:8-17
20. Albert MV, Kording K, Herrmann M, Jayaraman A. Fall classification by machine learning using mobile phones. *PLoS ONE*. 2012;7:e36556
21. Foster KR, Robert Koprowski, and Joseph D. Skufca. Machine learning, medical diagnosis, and biomedical engineering research-commentary. *Biomed. Eng. Online*. 2014;13:94
22. Pillow JW, Shlens, J., Paninski, L., Sher, A., Litke, A. M., Chichilnisky, E. J., Simoncelli, E. P. . Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature*. 2008;454:995
23. Benjamin AS, Fernandes, H. L., Tomlinson, T., Ramkumar, P., VerSteeg, C., Miller, L., Kording, K. P. . Modern machine learning far outperforms glms at predicting spikes. *bioRxiv*. 2017:111450
24. Hardoon DR, Sandor Szedmak, John Shawe-Taylor. Canonical correlation analysis: An overview with application to learning methods. *Neural Comput.* 2004;16:2639-2664
25. Andrew G, Arora, R., Bilmes, J., Livescu, K. . Deep canonical correlation analysis. *International Conference on Machine Learning*. 2013:1247-1255

26. Stevenson IH, Kording KP. How advances in neural recording affect data analysis. *Nat. Neurosci.* 2011;14:139-142
27. Glaser JI, Kording KP. The development and analysis of integrated neuroscience data. *Front Comput Neurosci*. 2016;10:11
28. O'Leary T, Alexander C. Sutton, Eve Marder. Computational models in the age of large datasets. *Curr. Opin. Neurobiol.* 2015;32:87-94
29. Gomez-Marin A. Causal circuit explanations of behavior: Are necessity and sufficiency necessary and sufficient? *Decoding neural circuit structure and function*. Cham: Springer; 2017:283-306.
30. Goodfellow I, Yoshua Bengio, Aaron Courville. *Deep learning.*: MIT Press; 2016.
31. Breiman L. Random forests. *Machine learning* 2001;45:5-32
32. Pedregosa F, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, et. al. . Scikit-learn: Machine learning in python. *The Journal of Machine Learning Research*. 2011;12:2825-2830
33. Dietterich TG. Ensemble methods in machine learning. *Multiple classifier systems* 2000;1857:1-15
34. Guyon I, Kristin Bennett, Gavin Cawley, Hugo Jair Escalante, Sergio Escalera, Tin Kam Ho, Núria Macia et al. . Design of the 2015 chalearn automl challenge. *Neural Networks (IJCNN), 2015 International Joint Conference on*. 2015:1-8
35. Corbett EA, Perreault EJ, Kording KP. Decoding with limited neural data: A mixture of time-warped trajectory models for directional reaches. *J. Neural Eng.* 2012;9:036002