How Can Physics Inform Deep Learning Methods in Scientific Problems?: Recent Progress and Future Prospects

Anuj Karpatne
University of Minnesota
karpa009@umn.edu

William Watkins, Jordan Read
United States Geological Survey
{wwatkins, jread}@umn.edu

Vipin Kumar
University of Minnesota
kumar001@umn.edu

Abstract

This work discusses a novel framework for learning deep learning models by using the scientific knowledge encoded in physics-based models. This framework, termed as physics-guided neural network (PGNN), leverages the output of physics-based model simulations along with observational features to generate predictions using a neural network architecture. Further, we discuss a novel class of learning objective for training neural networks, which ensures that the model predictions not only show lower errors on the training data but are also *consistent* with the known physics. We illustrate the effectiveness of PGNN for the problem of lake temperature modeling, where physical relationships between the temperature, density, and depth of water are used in the learning of neural network model parameters. By using scientific knowledge to guide the construction and learning of neural networks, we are able to show that the proposed framework ensures better generalizability as well as physical consistency of results.

1 Introduction

The growing deluge of data [2, 4, 10] has made long-lasting impacts on the way we sense, communicate, and make decisions in every walk of our life [8], through recent advances in data science methodologies such as deep learning. Apart from transforming commercial industries such as retail and advertising, deep learning is also beginning to play an important role in advancing scientific discovery. Historically, science has progressed by first generating hypotheses (or theories) and then collecting data to confirm or refute these hypotheses. However, in the big data era, ample data, which is being continuously collected without a specific theory or hypothesis in mind, offers further opportunity for discovering new knowledge. Based on the success of data science in applications where Internet-scale data is available (with billions or even trillions of samples), e.g., natural language translation, optical character recognition, object tracking, and most recently, autonomous driving, there is a growing anticipation of similar accomplishments in scientific disciplines [6, 11, 18]. To capture this excitement, some have even referred to the rise of data science in scientific disciplines as "the end of theory" [1], the idea being that the increasingly large amounts of data makes it possible to build actionable models without using scientific theories.

Unfortunately, this notion of black-box application of data science has met with limited success in scientific domains (e.g., [3, 15, 16]), for two main reasons. First, scientific problems are often under-constrained in nature as they suffer from paucity of representative training samples while involving a large number of physical variables. Further, physical variables commonly show complex and non-stationary patterns that dynamically change over time. For this reason, the limited number of labeled instances available for training or cross-validation can often fail to represent the true nature of relationships in scientific problems, leading to misleading conclusions. The paucity of representative samples is one of the prime challenges that differentiates scientific problems from mainstream

Workshop on Deep Learning for Physical Sciences (DLPS 2017), NIPS 2017, Long Beach, CA, USA.

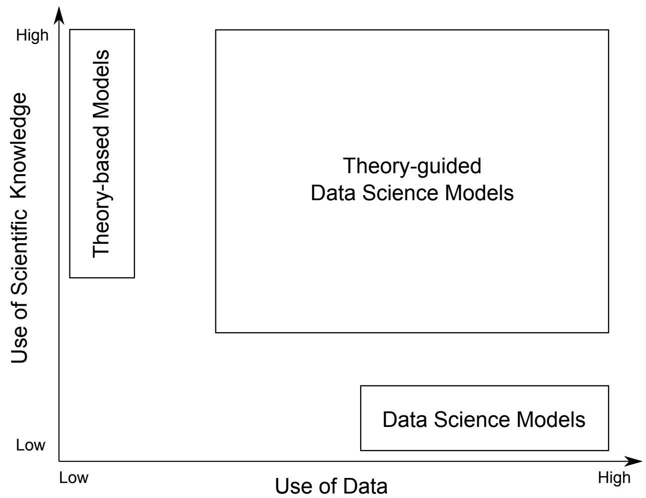

Figure 1: A representation of knowledge discovery methods in scientific applications. The x-axis measures the use of data while the y-axis measures the use of scientific knowledge. Theory-guided data science explores the space of knowledge discovery that makes ample use of the available data while being observant of the underlying scientific knowledge.

problems involving Internet-scale data such as language translation or object recognition, where large volumes of labeled or unlabeled data have been critical in the success of recent advancements in data science such as deep learning. Second, while a common end-goal in data science is the generation of actionable models, the process of knowledge discovery in scientific domains does not end at that. Rather, it is the translation of learned patterns and relationships to *interpretable* theories and hypotheses that leads to advancement of scientific knowledge, e.g., by explaining or discovering the physical cause-effect mechanisms between variables. Hence, even if a black-box model achieves somewhat more accurate performance but lacks the ability to deliver a mechanistic understanding of the underlying processes, it cannot be used as a basis for subsequent scientific developments. Further, an interpretable model, that is grounded by explainable theories, stands a better chance at safeguarding against the learning of spurious patterns from the data that lead to non-generalizable performance. This is especially important when dealing with problems that are critical in nature and associated with high risks (e.g., healthcare).

An alternate approach to black-box models for knowledge discovery is theory-based models, which are founded on core scientific principles and strive to advance our understanding of the physical world by learning explainable relationships between input and output variables. These models can range from solving closed-form equations (e.g. using Navier–Stokes equation for studying laminar flow) to running computational simulations of dynamical systems (e.g. the use of numerical models in climate science, hydrology, and turbulence modeling). However, a number of theory-based models use parameterized forms of approximations for representing complex physical processes that are either not fully understood or cannot be solved using computationally tractable methods. Calibrating the parameters in theory-based models is a challenging task because of the combinatorial nature of the search space. In particular, this can result in the learning of over-complex models that lead to incorrect insights even if they appear interpretable at a first glance. For example, these and other challenges in modeling hydrological processes using state-of-the-art theory-based models were the subject of a series of debate papers in Water Resources Research (WRR) [7, 14, 17].

Hence, neither a data-only nor a theory-only approach can be considered sufficient for knowledge discovery in complex scientific applications. Instead, there is a need to explore the continuum between theory-based and data science models, where both scientific theory and data are used in a synergistic manner. This is the paradigm of *theory-guided data science* (TGDS) [12] that attempts to address the shortcomings of data-only and theory-only models by seamlessly blending scientific knowledge in data science models (see Figure 1). By integrating scientific knowledge in data science models, TGDS aims to learn patterns and models that have a sufficient grounding in physical principles and thus have a better chance to represent causative relationships. TGDS further attempts to achieve better generalizability than models based purely on data by learning models that are consistent with scientific principles, termed as *physically consistent models*.

In this work, we discuss a novel framework that combines the power of deep learning with physics-based models, termed as *physics-guided neural networks* (PGNN) [13]. Specifically, we discuss a novel class of physics-based learning objective for training neural networks, which ensures that the learned networks not only admit to lower errors on the training data set but also produce outputs that are consistent with our scientific understanding of the physical world.

2 Physics-guided Neural Network

The framework of physics-guided neural networks (PGNN) [13] aims to integrate knowledge of physics in deep learning methods, to produce physically consistent outputs of neural networks. To illustrate the role of physical consistency in ensuring better generalization performance, consider the example of learning a neural network for a predictive learning problem using a limited supply of labeled samples. Ideally, we would like to learn a network that shows the best generalization performance over any unseen instance. Unfortunately, we can only observe the performance of a network on the available training set, which may not be truly representative of the true generalization performance (especially when the training size is small). In recognition of this fact, a number of learning frameworks have been explored to favor the selection of *simpler* models that may have lower accuracy on the training data (compared to more complex models) but are likely to have better generalization performance. This methodology, that builds on the well-known statistical principle of bias-variance trade-off [5], can be described using Figure 2.

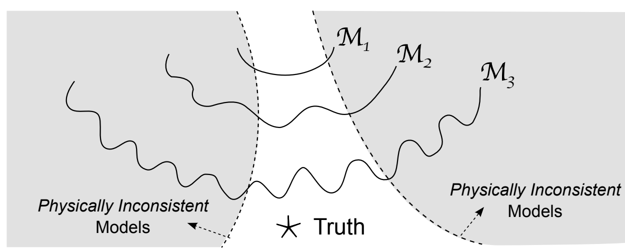

Figure 2: Scientific knowledge can help in reducing the model variance by removing physically inconsistent solutions, without likely affecting their bias.

Figure 2 shows an abstract representation of a succession of neural network families with varying levels of complexity (shown as curved lines), where $\mathcal{M}_1$ represents the set of least complex networks (with small number of hidden nodes) while $\mathcal{M}_3$ contains highly complex networks (with large of hidden nodes). Every point on the curved lines represents a neural network that a learning algorithm can arrive at, given a particular realization of training instances. The *true* relationship between the input and output variables is depicted as a star in Figure 2. We can observe that the learned models belonging to $\mathcal{M}_3$, on average, are quite close to the true relationship. However, even a small change in the training set can bring about large changes in the learned models of $\mathcal{M}_3$. Hence, $\mathcal{M}_3$ shows low *bias* but high *variance*. On the other hand, models belonging to $\mathcal{M}_1$ are quite robust to changes in the training set and thus show low variance. However, $\mathcal{M}_1$ shows high bias as its models are generally farther away from the true relationship as compared to models of $\mathcal{M}_3$. It is the trade-off between reducing bias and variance that is at the heart of a number of machine learning algorithms [19, 5, 20].

In scientific applications, there is another source of information that can be used to ensure the selection of generalizable models, which is the available scientific knowledge. By pruning candidate neural networks that are inconsistent with known scientific principles (shown as shaded regions in Figure 2), we can significantly reduce the variance of models without likely affecting their bias. A learning algorithm can then be focused on the space of physically consistent models, leading to generalizable and scientifically interpretable models. Hence, one of the overarching visions of TGDS is to include physical *consistency* as a critical component of model performance along with training accuracy and model complexity. This can be summarized in a simple way by the following revised objective of model performance in TGDS:

$$\text{Performance} \propto \text{Accuracy} + \text{Simplicity} + Consistency.$$

There are various ways of introducing physical consistency in deep learning methods, in different forms and capacities. First, scientific knowledge can be used in the design of neural network architecture to restrict the space of models to physically consistent solutions, e.g., in the selection of activation function or the pattern of connections among the layers. Second, we can also guide a neural network learning algorithm to focus on physically consistent solutions, e.g., by initializing the model with physically meaningful parameters, by encoding scientific knowledge as probabilistic relationships, by using domain-guided constraints, or with the help of regularization terms or loss functions inspired by our physical understanding. Third, the outputs of a neural network can also be refined using available scientific knowledge, by using pruning or post-processing methods.

In our recent work [13], we demonstrate an application of PGNN for modeling the temperature of water in lakes. For this problem, a number of physics-based models have been developed that involve parameters (e.g., parameters related to vertical mixing, wind energy inputs, and water clarity) whose values can be set to default values or custom-calibrated for each lake if some training data is available. Because this step of custom-calibrating is both labor- and computation-intensive, there is a trade-off between increasing the accuracy of the model and expanding the feasability of study to a large number of lakes. In our work, we used the simulation outputs of a physics-based model, namely General Lake Model (GLM) [9], as input variables in the neural network framework to construct hybrid-physics-data models. We further explored a novel class of physics-based loss functions that exploit the physical relationships between the temperature estimates produced by the neural network with other physical variables such as density of water at varying depths in the lake. The resultant PGNN model shows improved generalization performance than both the black-box neural network model (that is agnostic to the underlying physics) as well the state-of-the-art physics-based model (that does not make effective use of the available data). The PGNN model additionally produces physically meaningful results that can be used as inputs in other models of lake properties such as water quality, thus resulting in advancement of scientific discovery.

References

[1] C. Anderson. The end of theory. *Wired magazine*, 16(7):16–07, 2008.

[2] G. Bell, T. Hey, and A. Szalay. Beyond the data deluge. *Science*, 323(5919):1297–1298, 2009.

[3] P. M. Caldwell, C. S. Bretherton, M. D. Zelinka, S. A. Klein, B. D. Santer, and B. M. Sanderson. Statistical significance of climate sensitivity predictors obtained by data mining. *Geophysical Research Letters*, 41(5):1803–1808, 2014.

[4] Economist. The data deluge. *Special Supplement*, 2010.

[5] J. Friedman, T. Hastie, and R. Tibshirani. *The elements of statistical learning*, volume 1. Springer series in statistics Springer, Berlin, 2001.

[6] D. Graham-Rowe, D. Goldston, C. Doctorow, M. Waldrop, C. Lynch, F. Frankel, R. Reid, S. Nelson, D. Howe, S. Rhee, et al. Big data: science in the petabyte era. *Nature*, 455(7209):8–9, 2008.

[7] H. V. Gupta and G. S. Nearing. Debates—the future of hydrological sciences: A (common) path forward? using models and data to learn: A systems theoretic perspective on the future of hydrological science. *Water Resources Research*, 50(6):5351–5359, 2014.

[8] A. Halevy, P. Norvig, and F. Pereira. The unreasonable effectiveness of data. *Intelligent Systems, IEEE*, 24(2):8–12, 2009.

[9] M. Hipsey, L. Bruce, and D. Hamilton. Glm—general lake model: Model overview and user information. *Perth (Australia): University of Western Australia Technical Manual*, 2014.

[10] M. James, C. Michael, B. Brad, B. Jacques, D. Richard, R. Charles, and H. Angela. Big data: The next frontier for innovation, competition, and productivity. *The McKinsey Global Institute*, 2011.

[11] T. Jonathan, A. Gerald, et al. Special issue: dealing with data. *Science*, 331(6018):639–806, 2011.

[12] A. Karpatne, G. Atluri, J. H. Faghmous, M. Steinbach, A. Banerjee, A. Ganguly, S. Shekhar, N. Samatova, and V. Kumar. Theory-guided data science: A new paradigm for scientific discovery from data. *IEEE Transactions on Knowledge and Data Engineering*, 29(10):2318–2331, 2017.

[13] A. Karpatne, W. Watkins, J. Read, and V. Kumar. Physics-guided neural networks (pgnn): An application in lake temperature modeling. *arXiv preprint arXiv:1710.11431*, 2017.

[14] U. Lall. Debates—the future of hydrological sciences: A (common) path forward? one water. one world. many climes. many souls. *Water Resources Research*, 50(6):5335–5341, 2014.

[15] D. Lazer, R. Kennedy, G. King, and A. Vespignani. The Parable of Google Flu: Traps in Big Data Analysis. *Science (New York, N.Y.)*, 343(6176):1203–5, Mar. 2014.

[16] G. Marcus and E. Davis. Eight (no, nine!) problems with big data. *The New York Times*, 6(04):2014, 2014.

[17] J. J. McDonnell and K. Beven. Debates—the future of hydrological sciences: A (common) path forward? a call to action aimed at understanding velocities, celerities and residence time distributions of the headwater hydrograph. *Water Resources Research*, 50(6):5342–5350, 2014.

[18] T. J. Sejnowski, P. S. Churchland, and J. A. Movshon. Putting big data to good use in neuroscience. *Nature neuroscience*, 17(11):1440–1441, 2014.

[19] P.-N. Tan, M. Steinbach, and V. Kumar. *Intorduction to Data Mining*. Addison-Wesley, 2005.

[20] V. N. Vapnik and V. Vapnik. *Statistical learning theory*, volume 1. Wiley New York, 1998.