CRUNCH Seminars at Brown, Division of Applied Mathematics

Friday - April 20, 2020

Theory, Algorithms, and Software for Physics-Informed Deep Learning

Lu Lu

This dissertation is centered around the broad topic of physics-informed deep learning. Specifically, it covers the following three fronts:

Algorithms Recent developments of deep learning (DL), i.e., deep neural networks (NNs), provide us with opportunities that cannot be tackled solely through traditional methods for scientific applications. However, DL usually requires a large amount of data, and in many science and engineering problems, it is often difficult to obtain the necessary data of high accuracy. To address this challenge, there is a significant opportunity for DL to complement and also augment significantly traditional scientific domain knowledge, including physical principles, constraints, symmetries, computational simulations, etc. By incorporating such scientific domain knowledge in DL, it is expected to improve the accuracy, interpretability, and robustness, while simultaneously reducing data requirements and accelerating NN training and prediction. Specifically, I have developed three new algorithms to incorporate physics into DL. 1. Multi-fidelity NNs. Instrumented indentation has been developed and widely utilized as one of the most versatile and practical means of extracting mechanical properties of materials. I present several multi-fidelity approaches for solving the inverse indentation problem, hence significantly reducing the number of expensive high-fidelity datasets required to achieve a given level of accuracy, and utilize known physical and scaling laws to improve training efficiency and accuracy. 2. Physics-informedNNs(PINNs). I further develop PINNs, which embed partial differential equations (PDEs) into NNs. The PINN algorithm is simple, and it can be applied to different types of PDEs, including integro-differential equations, fractional PDEs, and stochastic PDEs. Moreover, from the implementation point of view, PINNs solve inverse problems as easily as forward problems. I propose a new residual-based adaptive refinement method to improve the training efficiency of PINNs. 3. DeepONet: Learning nonlinear operators. While it is widely known that NNs are universal approximators of continuous functions, a less known but more powerful result is that a NN with a single hidden layer can approximate accurately any nonlinear continuous operator. To realize this theorem in practice, I propose the deep operator network (DeepONet) to learn operators accurately and efficiently from a relatively small dataset. I demonstrate that DeepONet significantly reduces the generalization error. More importantly, I observe high-order error convergence through computational tests, and even exponential convergence with respect to the training dataset size.

Theory Building the mathematical, statistical, and information-theoretic foundations of DL is especially vital to establishing assurance and thus realizing the potential of DL for science and engineering. However, despite the remarkable success in the last 15 years, our theoretical understanding of DL is lagging behind, which is currently a serious bottleneck of DL for scientific discovery.
I theoretically address the question of assurance: whether a DL model is appropriately trained and used for the task for which it is intended, including whether prediction errors can be meaningfully bounded. 1. Dying ReLU. The dying ReLUrefers to the problem when ReLUneurons become inactive and only output 0 for any input. There are many empirical and heuristic explanations of why ReLU neurons die. However, little is known about its theoretical analysis. I rigorously prove that a deep ReLU network will eventually die in probability as the depth goes to infinite. I also propose a new initialization procedure, namely, a randomized asymmetric initialization. I show that the new initialization can effectively prevent the dying ReLU. 2. Generalization. The accuracy of DL can be characterized by dividing the total error into three main types: approximation, optimization and generalization. Whereas there are some satisfactory answers to approximation and optimization errors, much less is known about the theory of generalization. I study the generalization error of NNs for classification problems in terms of data distribution and NN smoothness. I introduce the cover complexity (CC) as a new metric to measure the difficulty of learning a data set and the inverse of the modulus of continuity to quantify NN smoothness. A quantitative bound for expected accuracy/error is derived by considering both the CC and NN smoothness.

Open-source software I also present a Python library for PINNs, DeepXDE, which is designed to serve both as an education tool to be used in the classroom as well as a research tool for solving general problems in computational science and engineering. Specifically, DeepXDE can solve forward problems given initial and boundary conditions, as well as inverse problems given some extra measurements. More broadly, DeepXDE contributes to the more rapid development of the emerging Scientific Machine Learning field.