Speakers

1 Mohammad Emtiyaz Khan

2 Jiequn Han

3 Nan Jiang

4 Nihar Shah

5 Greg Yang

Abstracts

Machine Learning and Its Applications

(10–28 Oct 2022)

1 Mohammad Emtiyaz Khan

RIKEN, Japan
Deep learning from Bayesian principles

Abstract

I will present the Bayesian learning rule, a new learning rule based on Bayesian principles, and show how we can derive deep-learning algorithms as special instances of it. We will use the rule to derive new Bayesian deep-learning algorithms which can estimate uncertainty without much computational overhead.

I will briefly discuss the importance of the new learning rule for adaptive AI systems.

2 Jiequn Han

Flatiron Institute, USA
Deep learning algorithms for partial differential equations

Abstract

In recent years, tremendous progress has been made on deep learning-based numerical algorithms for solving partial differential equations (PDEs) in high dimensions. They are potentially free of the curse of dimensionality for many different applications. This lecture will review these numerical advances. The primary examples are the Deep BSDE method and variational Monte

Carlo method, the first two algorithms in the literature for high-dimensional problems using modern deep learning. In addition, we also discuss deep learning algorithms based on the traditional Ritz, Galerkin, and least square formulations (physics-informed neural network). We hope to demonstrate to the audience that studying PDEs as well as control and variational problems in very high dimensions might very well be among the most promising new directions in mathematics and scientific computing in the near future.

3 Nan Jiang

University of Illinois Urbana-Champaign, USA
Offline RL theory

Abstract

Reinforcement learning (RL) studies sequential decision-making in unknown environments and has seen fast theoretical advances in the last few years, especially in the realistic regime of large state (and action) spaces that require (possibly nonlinear) function approximation. In this tutorial, I will use offline RL (i.e., evaluating/learning policies from historical data without online interactions) as an example setting to introduce the theoretical foundations of RL under function approximation, while also making connections to online RL. I will describe important concepts in RL theory (e.g., realizability, Bellman-completeness, density ratio, data coverage, etc.), the major assumptions that enable the theoretical guarantees, the representative algorithms, and the key ideas behind the theoretical analyses.

4 Nihar Shah

Carnegie Mellon University, USA
AI-augmented human evaluations

Abstract

Many important applications such as university admissions, peer review, hiring, healthcare, judicial decisions etc. involve evaluations of a large number of items by a set of evaluators. These evaluations are typically distributed – each evaluator evaluates a small number of items and each item is evaluated by a small number of evaluators. This distributed nature of evaluations

leads to a number of problems including fraud, biases, inefficiencies, miscalibration, noise, subjectivity, etc. We will discuss the design and use of AI techniques to augment human evaluations and mitigate these problems.

A survey article (focused on the application of peer review) is available here https://www.cs.cmu.edu/~nihars/preprints/SurveyPeerReview.pdf

5 Greg Yang

Microsoft Research, USA

The unreasonable effectiveness of mathematics in large scale deep learning

Abstract

Recently, the theory of infinite-width neural networks led to the first technology, muTransfer, for tuning enormous neural networks that are too expensive to train more than once. For example, this allowed us to tune the 6.7 billion parameter version of GPT-3 using only 7% of its pretraining compute budget, and with some asterisks, we get a performance comparable to the original GPT-3 model with twice the parameter count. In this talk, I will explain the core insight behind this theory. In fact, this is an instance of what I call the *Optimal Scaling Thesis*, which connects infinite-size limits for general notions of "size" to the optimal design of large models in practice, illustrating a way for theory to reliably guide the future of AI. I'll end with several concrete key mathematical research questions whose resolutions will have incredible impact on how practitioners scale up their NNs.