



Events

[Current Events](#)[Lectures](#)[Events Archive](#)

Other - Colloquium on Artificial Intelligence Research and Optimization

SLIDE: Commodity Hardware is All You Need for Large-Scale Deep Learning**Anshumali Shrivastava, Rice University**

Assistant Professor

Virtual- REGISTRATION REQUIRED (SEE ABSTRACT) Zoom
March 17, 2021 - 01:00 pm**Abstract:**

Current Deep Learning (DL) architectures are growing larger to learn from complex datasets. The trends show that the only sure-shot way of surpassing prior accuracy is to increase the model size, supplement it with more data, followed by aggressive fine-tuning. However, training and tuning astronomical sized models are time-consuming and stall the progress in AI. As a result, industries are increasingly investing in specialized hardware and deep learning accelerators like GPUs to scale up the process. It is taken for granted that commodity hardware CPU is incapable of outperforming powerful accelerators such as V100 GPUs in a head-to-head comparison of training large DL models. However, GPUs come with additional concerns: expensive infrastructural change, hard to virtualize, main memory limitations.

In this talk, I will demonstrate the first algorithmic progress that challenges the common knowledge prevailing in the community that specialized processors like GPUs are significantly superior to CPUs for training large neural networks. The algorithm is a novel alternative to traditional matrix-multiplication-based backpropagation. We will show how data structures, particularly hash tables, can reduce the no of multiplications associated with the forward pass of the neural networks. The very sparse nature of updates uniquely allows for an asynchronous data-parallel gradient descent algorithm. A C++ implementation with multi-core parallelism and workload optimization on CPU is anywhere from 4-15x faster than the most optimized implementations of Tensorflow on the best available V100 GPUs in a head to head comparisons. The associated task is training a 200-million-parameter neural network on Kaggle Amazon recommendation datasets.

REGISTRATION IS FREE AND REQUIRED TO RECIEVE ZOOM ID:

<https://docs.google.com/forms/d/e/1FAIpQLSdLclBvX8tiTkxm9HbvBdyPx70fhpkpzsAQmCf0AXgn44cqcQ/viewform>
Speaker's Bio:

Anshumali Shrivastava is an assistant professor in the computer science department at Rice University. His broad research interests include randomized algorithms for large-scale machine learning. In 2018, Science news named him one of the Top-10 scientists under 40 to watch. He is a recipient of the National Science Foundation CAREER Award, a Young Investigator Award from the Air Force Office of Scientific Research, and a machine learning research award from Amazon. He has won numerous paper awards, including Best Paper Award at NIPS 2014 and Most Reproducible Paper Award at SIGMOD 2019. IEEE Spectrum describes his work on scaling up deep learning as,

"stunning." Investorplace considers SLIDE algorithm one of the biggest threats to NVIDIA Stock.

