Nima Shoghi

https://nima.sh · nima@nima.sh

Objective: Dedicated AI researcher with a strong engineering foundation and practical AI project experience. Seeking research roles in preparation for a Ph.D.

Education

Georgia Institute of Technology — Atlanta, GA

MS Computer Science (Machine Learning Specialization) — *Highest Honors* — Jan 2020 - May 2021

BS Computer Science (Machine Learning and Devices Threads) — *High Honors* — May 2015 - Dec 2019

Selected Coursework: Design & Analysis of Algorithms, Automata & Complexity, Robotics & Perception, Computer Vision, Machine Learning, Markov Chains & Emergence, Convex Optimization, Deep Learning, DL for Text Data

Work Experience

Meta Fundamental AI Research (FAIR) — Menlo Park, CA and Remote

AI Resident on the FAIR Chemistry Team — Aug 2021 - Aug 2023

Worked on the Open Catalyst Project, which aims to use AI to discover new catalysts for use in renewable energy storage to help in addressing climate change. Tasks include:

- Developed large foundation models for atomic property prediction, pre-trained on data from diverse chemical domains. Fine-tuned the model to achieve state-of-the-art results across 30/40 tasks on the rMD17, QM9, MatBench, QMOF, SPICE, and MD22 datasets and competitive results on the remaining 10 tasks.
- Designed and implemented novel attention-based spatiotemporal graph neural networks for predicting force fields in relaxation trajectories and molecular dynamics simulations
- Contributed to the creation of the Open Catalyst 2022 dataset, featuring a comprehensive set of 62K Density Functional Theory relaxations across a range of oxide materials, coverages, and adsorbates. Evaluated the performance of state-of-the-art models on this dataset as a benchmark for future work.
- Led the transformation of the team's codebase to a more modular and extensible codebase, with a focus on reproducibility and ease of use. Added out-of-the-box support for multi-domain supervised pre-training, used for the development of large foundation models used internally on the FAIR Chemistry team.

High Performance Computer Architecture Lab at Georgia Institute of Technology — Atlanta, GA

Graduate Research Assistant — May 2019 - May 2021

Developed software-level and hardware-level techniques for accelerating deep learning training and inference. Co-authored 7 papers (3 first-author) published in various conferences and journals.

Ciena Corporation — Atlanta, GA

Software Engineering Intern — May 2017 - May 2018

Developed software to interface with network devices for orchestration, which was deployed to over 150 customers worldwide. Maintained build tools and CI/CD pipeline for the application build process. Created tools for measuring and optimizing developer productivity.

Professional Skills

Data Science and Machine Learning

- Proficient in Python data science libraries: NumPy, Pandas, Matplotlib, and Seaborn.
- Extensive experience with deep learning libraries: PyTorch and JAX.
- Experience with high-performance computing (HPC) and distributed (e.g., 128+ GPUs) training.

Programming and Development

- Proficient in Python, C, C++, Rust, C#, and JavaScript/TypeScript
- Experience with test-driven development, including unit tests, integration tests, and end-to-end tests.
- In-depth knowledge of virtualization, containers, and Docker.

Publications (* denotes equal contribution)

Context-Aware Task Handling in Resource-Constrained Robots with Virtualization

*Ramyad Hadidi, **Nima Shoghi**, Bahar Asgari, Hyesoon Kim*

IEEE International Conference on Edge Computing and Communications, EDGE 2023, Chicago, IL, USA, July 2-8, 2023, 2023

Presents a new context-aware approach for handling tasks in real-time on resource-constrained robots, achieving increased execution speed by integrating a dynamic time-sharing mechanism, event-driven scheduling, and lightweight virtualization.

The Open Catalyst 2022 (OC22) dataset and challenges for oxide electrocatalysts

*Richard Tran, Janice Lan, Muhammed Shuaibi, Brandon Wood, Siddharth Goyal, Abhishek Das, Javier Heras-Domingo, Adeesh Kolluru, Ammar Rizvi, **Nima Shoghi**, Anuroop Sriram, Félix Therrien, Jehad Abed, Oleksandr Voznyy, Edward Sargent, Zachary Ulissi, C. Zitnick*

ACS Catalysis, 2023

Introduces the open source OC22 dataset, containing relaxations across various oxide materials, coverages, and adsorbates, to improve machine learning models for oxide electrocatalysts. The paper also establishes clear benchmarks for future efforts in this area, opens the data and models for community development, and introduces a public leaderboard.

Open Challenges in Developing Generalizable Large-Scale Machine-Learning Models for Catalyst Discovery
Adeesh Kolluru, Muhammed Shuaibi, Aini Palizhati, ***Nima Shoghi****, Abhishek Das, Brandon Wood, C Zitnick, John Kitchin, Zachary Ulissi*
ACS Catalysis, 2022

Examines the challenges and limited generalizability of current machine-learning models for catalyst discovery and discusses the potential advancements brought about by large-scale catalyst data sets like OC20.

Transfer learning using attentions across atomic systems with graph neural networks (TAAG)
Adeesh Kolluru, ***Nima Shoghi****, Muhammed Shuaibi, Siddharth Goyal, Abhishek Das, C Zitnick, Zachary Ulissi*
The Journal of Chemical Physics, 2022

Introduces TAAG, a novel attention-based transfer learning approach for Graph Neural Networks which significantly improves performance for out-of-domain datasets and speeds up model training, demonstrating the potential for generalizing important aspects across different atomic system domains.

SmaQ: Smart Quantization for DNN Training by Exploiting Value Clustering
Nima Shoghi*, Andrei Bersatti*, Moinuddin Qureshi, Hyesoon Kim*
IEEE Computer Architecture Letters, 2021

Introduces Smart Quantization (SmaQ), a quantization scheme that leverages the normal distribution properties of neural network data structures, leading to a memory usage reduction of up to 6.7x during training, with minimal impact on accuracy.

Quantifying the design-space tradeoffs in autonomous drones
Ramyad Hadidi, Bahar Asgari, Sam Jijina, Adriana Amyette, ***Nima Shoghi****, Hyesoon Kim*
Proceedings of the 26th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, 2021

Examines the inherent design complexities in autonomous drones—especially the trade-offs between compute, energy, and electromechanical resources—and proposes a systematic exploration of the drone design space. The study emphasizes the benefits of optimizing the SLAM process on FPGA platforms and introduces a customizable, open-source drone.

Secure Location-Aware Authentication and Communication for Intelligent Transportation Systems
Nima Shoghi*, Ramyad Hadidi, Lee Jaewon, Jun Chen, Arthur Siqueria, Rahul Rajan, Shaan Dhawan, Pooya Shoghi, Hyesoon Kim*
arXiv preprint arXiv:2011.08936, 2020

Introduces a unique location-aware protocol for secure communication in Intelligent Transportation Systems, leveraging in situ visual localization (like QR codes) for efficient message verification. The new method is efficient, scalable, and infrastructure-independent, providing a more trustworthy and widely applicable solution than previous approaches.

Understanding the Software and Hardware Stacks of a General-Purpose Cognitive Drone
Sam Jijina, Adriana Amyette, ***Nima Shoghi****, Ramyad Hadidi, Hyesoon Kim*
2020 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS), 2020

Conducts a detailed analysis of drone operation and efficiency by exploring hardware and software components, using ArduCopter as an example. Optimizing specific aspects of these components can significantly increase drone flight range.

PISCES: Power-Aware Implementation of SLAM by Customizing Efficient Sparse Algebra
Bahar Asgari, Ramyad Hadidi, ***Nima Shoghi****, Hyesoon Kim*
2020 57th ACM/IEEE Design Automation Conference (DAC), 2020

Introduces Pisces, a method that optimizes power consumption and latency for simultaneous localization and mapping (SLAM). Through using sparse data and reducing memory access, it results in a 2.5 times power reduction and 7.4 times faster execution than other contemporary methods.

Neural Network Weight Compression with NNW-BDI
Andrei Bersatti, **Nima Shoghi***, Hyesoon Kim*

The International Symposium on Memory Systems, 2020

Introduces NNW-BDI, a specialized memory compression scheme for neural network weights, successfully decreasing memory usage by up to 85

SLAM Performance on Embedded Robots
***Nima Shoghi**, Ramyad Hadidi, Hyesoon Kim*
Student Research Competition at Embedded System Week (SRC ESWEEK), 2019

Examines the effectiveness of the ORBSLAM2 algorithm on the Raspberry Pi for real-time usage in embedded robots, identifying the Pi's slower performance, but proposing optimizations that nearly quintuple its speed with minimal precision loss, enabling real-time operation.