

Neil Adit

PhD student, Computer Systems Lab, Cornell University

☎ (+1) 646-620-1151 | ✉ na469@cornell.edu | 🏠 neiladit.com | 📄 github.com/neiladit

I am a research scientist at Meta working on developing efficient foundational recommendation systems at scale. Previously, I was a student researcher at Google and PhD candidate at Cornell, working on hardware-software co-design for datacenter workloads. My research interests lie at the intersection of Efficient Machine Learning, Compilers and Computer Architecture.

Education

Cornell University

Aug 2018 - June 2024

PHD IN COMPUTER ENGINEERING

GPA: 4/4

- Committee: Adrian Sampson (chair), Zhiru Zhang, Chris De Sa
- Courses: Advanced ML Systems, Computer Vision, Advanced Compilers, Datacenter Computing, Parallel Computing

Indian Institute of Technology, Bombay

July 2013 - June 2018

B.TECH + M.TECH IN ELECTRICAL ENGINEERING

GPA: 9.05/10

- Masters in Microelectronics, Minor in Computer Science
- Advisor: Sachin Patkar

Publications

Performance Left on the Table: An Evaluation of Compiler Auto-Vectorization for RISC-V

IEEE Micro 2022

Neil Adit and Adrian Sampson

Software-Defined Vector Processing on Manycore Fabrics

MICRO 2021

Philip Bedoukian, Neil Adit, Edwin Peguero, Adrian Sampson

Dense Pruning of Pointwise Convolutions in the Frequency Domain

arxiv preprint 2021

Mark Buckler, Neil Adit, Yuwei Hu, Zhiru Zhang, and Adrian Sampson

Dagger: Efficient and Fast RPCs in Cloud Microservices with Near-Memory Reconfigurable NICs

ASPLOS 2021

Nikita Lazarev, Shaojie Xiang, Neil Adit, Zhiru Zhang, Christina Delimitrou

Dagger: Towards Efficient RPCs in Cloud Microservices with Near-Memory Reconfigurable NICs

IEEE CAL 2020

Nikita Lazarev, Neil Adit, Shaojie Xiang, Zhiru Zhang, and Christina Delimitrou

Industry Experience

Meta

Menlo Park, CA, USA

RESEARCH SCIENTIST

July 2024 - Present

- Developing efficient large recommendation system models, optimized for various hardware backends

Google

Sunnyvale, CA, USA

STUDENT RESEARCHER | ADVISORS: AKANKSHA JAIN AND SNEHASISH KUMAR

May 2022 - May 2024

- Developed a hardware-software co-design infrastructure for GSoC performance optimization, including profile-driven compiler analysis
- Demonstrated performance improvements via microarchitectural modifications in cycle-accurate simulator, modeling datacenter behaviour

Microsoft Research

Redmond, WA, USA

RESEARCH INTERN | ADVISOR: OFER DEKEL

May 2021 - Aug 2021

- Developed algorithms to accelerate sparse ML models on commodity hardware in the Machine Learning and Optimization group, at MSR
- Demonstrated wall-clock speedups on sparse kernel execution using the ONNX runtime library backend

Intel Labs

Santa Clara, CA, USA

GRADUATE RESEARCH INTERN | ADVISOR: FABRIZIO PETRINI

May 2019 - Aug 2019

- Designed and implemented high performance computing algorithms for sparse computations on Intel's breakthrough research architecture

SIEMENS Research

Bangalore, India

SUMMER INTERN | ADVISORS: DR. AMIT KALE AND PRABHU TEJA

May 2016 - Jul. 2016

- Designed and demonstrated Kidney segmentation in CT images for clinical diagnosis using Laplacian Mesh Deformation

Research Experience

Compiler Auto-Vectorization for Scalable Vectors

Ithaca, NY, USA

ADVISOR: PROF. ADRIAN SAMPSON, CORNELL UNIVERSITY

Aug. 2021 - May 2022

- Identified compiler limitations in front-end (auto-vectorization pass), IR and backend (instruction selection), for length agnostic vector ISAs
- Designed backend-independent ScaleIR for arbitrary representations to optimize instruction selection and hardware performance

Software-Defined Vectors on Manycore

Ithaca, NY, USA

ADVISOR: PROF. ADRIAN SAMPSON, CORNELL UNIVERSITY

Jan. 2019 - Aug. 2020

- Worked with Philip Bedoukian on vector programming model that allows dynamic reconfiguration of manycore tiles into vector engines

Frequency Domain Dense Pruning of Pointwise Convolutions

Ithaca, NY, USA

ADVISOR: PROF. ADRIAN SAMPSON, CORNELL UNIVERSITY

Aug. 2018 - Sep. 2021

- Worked with Mark Buckler on exploiting spatial redundancy in depthwise convolutions by pruning in the frequency domain

Near-Memory Reconfigurable NICs

Ithaca, NY, USA

ADVISOR: PROF. CHRISTINA DELIMITROU, CORNELL UNIVERSITY

Jan. 2020 - Aug. 2020

- We offload the RPC stack on a FPGA which is tightly coupled with the host CPU via memory interconnects, Intel UPI in this case
- Designed queue management for asynchronously sending packets in a single connection

Accelerating 1x1 Convolutions using Systolic Arrays

Ithaca, NY, USA

ADVISOR: PROF. ZHIRU ZHANG, CORNELL UNIVERSITY

Oct. 2018 - Dec. 2018

- Implemented pointwise convolutions in MobileNets on Zynq ZC-706 using systolic arrays.
- Optimized streaming of input activations using quantization, bit packing and padding.
- Achieved close to ideal, 215x speedup using 16x16 parallel PEs for systolic array architecture.

Parallel Sparse Matrix Solution on FPGA

Mumbai, India

ADVISOR: PROF. SACHIN PATKAR, IIT BOMBAY

Jul. 2017 - Jun. 2018

- Accelerating sparse matrix solvers for performance improvements in SPICE circuit simulators
- Designing a stack based processor with pipelined FPU to process LU expressions parallelly
- Achieved upto 6x speedup using parallel hardware directives, optimizing off-chip memory access and minimizing arithmetic operations

Person Re-Identification using Deep Learning

Mumbai, India

ADVISOR: PROF. SUBHASIS CHAUDHURI, IIT BOMBAY

Jul. 2017 - Dec. 2017

- Developing a Deep Learning model to spot person of interest across cameras for surveillance applications
- Modelled a RNN (temporal features) and CNN (spatial features) based Siamese network for video-based re-identification in Torch
- Achieved rank-1 accuracy 60% comparable to state-of-the-art with smaller test image sequence on iLIDS-VID dataset

Academic Achievements

2018	Eastman Fellowship , Cornell University	Ithaca, U.S.A
2017	Excellence in Teaching Assistantship , EE, IIT Bombay	Mumbai, India
2013	All India Rank 242 , IIT Joint Entrance Exam (JEE)-Advanced, among 1.4 million examinees	India
2012	Ranked 115 , KVPY Scholarship, Govt. of India , among 200,000 candidates	India
2012	Top 1% , National Physics Olympiad	Delhi, India
2011	Ranked 20 , Regional Mathematics Olympiad (RMO) and among top 900 nationally to compete in Indian National Mathematical Olympiad (INMO)	Delhi, India

Extracurricular Activity

Institute Student Mentorship Programme (ISMP)

IIT Bombay, India

STUDENT MENTOR

Aug. 2016 - May. 2018

Selected for 2 consecutive years as part of team of 82 mentors from 368 applicant.

- Mentored 24 students for smooth transition to campus life, supporting their academic & co-curricular endeavors

Formula Student Racing Team

IIT Bombay, India

DESIGN ENGINEER

Sep. 2014 - Apr. 2016

A team of 70 students building India's fastest electric racecar for Formula Student UK, an international race car design competition. Won FS Award for 2 consecutive years worth £3000 (2 out of 48 non-UK teams) for major design improvements

- Headed a team of 5 Engineers to design onboard data logging and real-time remote wireless data monitoring GUI system
- Implemented team's first CAN protocol to improve reliability in communication and reduce wiring harness
- Implemented Electronic Differential and Regenerative Braking for the first time within the team
- Achieved 2x reduction in size and weight of PCB enclosure by optimized routing in boards