Supun Nakandala

Phone: (+1) 812-558-6888
Email: snakanda@eng.ucsd.edu
Web: https://scnakandala.github.io

3232 EBU3B CSE
9500 Gilman Drive
La Jolla, CA 92093

Research Interests My research interest lies broadly in the intersection of Systems and Machine Learning, an emerging area which is increasingly referred to as *Machine Learning Systems*. In this space I operate as a data management researcher. My Ph.D. thesis work focuses on developing query optimization-inspired abstractions, algorithms, and systems to improve efficiency, scalability, and usability of deep learning workloads.

Education

University of California, San Diego — Sept 2017 - (expected) March 2022
PhD, Computer Science.
Thesis Title: *Multi-Query Optimization for Deep Learning Systems*
Thesis Advisor: Arun Kumar
Committee Members: Arun Kumar (chair), Yannis Papakonstantinou, Geoffrey Voelker, Lawrence Saul, Loki Natarajan

University of California, San Diego — Sept 2017 - June 2020
MSc, Computer Science. GPA: 3.97/4.00

University of Moratuwa, Sri Lanka — Aug 2010 - April 2015
Bachelor of the Science of Engineering, Computer Science & Engineering.
GPA: 4.11/4.20.
Department Topper and Gold Medalist

Work Experience

Research Intern — June 2021 - Sept 2021
Data Management, Exploration, and Mining Group – Microsoft Research
(Incoming Intern)

Software Engineering Intern — June 2020 - Sept 2020
Redshift – Amazon Web Services
Mentor: Yannis Papakonstantinou
Designed and implemented components of Redshift ML, in-database ML feature of Redshift data warehouse, that went to public preview.

Research Intern — June 2019 - Sept 2019
Gray Systems Lab – Microsoft
Mentors: Matteo Interlandi, Markus Weimer
Designed and implemented Hummingbird system, a compiler for translating classical machine learning pipelines into tensor computations for unified and faster scoring of machine learning models.

Research Software Developer — Oct 2015 - Aug 2017
Science Gateways Research Center, Indiana University
Manager: Marlon Pierce
Contributed to the development of Apache Airavata system, which is a software framework to compose, manage, execute, and monitor computational applications and workflows on distributed computing resources such as local clusters, computational grids, and computing clouds.

Teaching Experience

Teaching Assistant - Systems for Scalable Analytics — UCSD - Winter 2020
Teaching Assistant - Advanced Data Analytics Systems — UCSD - Spring 2019

Conference Publications

Cerebro: A Layered Data Platform for Scalable Deep Learning
Arun Kumar, **Supun Nakandala**, Yuhao Zhang, Side Li, Advitya Gemawat, and Kabir Nagrecha
CIDR 2021 (Vision paper)

Cerebro: A Data System for Optimized Deep Learning Model Selection
Supun Nakandala, Yuhao Zhang, and Arun Kumar
VLDB 2020

A Tensor Compiler for Unified Machine Learning Prediction Serving
Supun Nakandala, Karla Saur, Gyeong-In Yu, Konstantinos Karanasos, Carlo Curino, Markus Weimer, and Matteo Interlandi
OSDI 2020

Vista: Declarative Feature Transfer from Deep CNNs at Scale
Supun Nakandala and Arun Kumar
SIGMOD 2020

Extending Relational Query Processing with ML Inference
Konstantinos Karanasos, Matteo Interlandi, Doris Xin, Fotis Psallidas, Rathijit Sen, Kwanghyun Park, Ivan Popivanov, **Supun Nakandala**, Subru Krishnan, Markus Weimer, Yuan Yu, Raghu Ramakrishnan, Carlo Curino
CIDR 2020

Incremental and Approximate Inference for Faster Occlusion-based Deep CNN Explanations
Supun Nakandala, Arun Kumar, and Yannis Papakonstantinou
SIGMOD 2019
Honorable Mention for Best Paper Award
Invited to TODS 2020
Invited to SIGMOD Research Highlight 2020

Gendered Conversation in a Social Game-Streaming Platform
Supun Nakandala, Giovani Cimpaglia, Norma Su, and Yong-Yeol Ahn
AAAI ICWSM 2017

Apache Airavata Security Manager: Authentication and Authorization Implementations for a Multi-Tenant eScience Framework
Supun Nakandala, Hasini Gunasinghe, Suresh Marru, and Marlon Pierce
IEEE e-Science 2016

Journal Publications

The CNN Hip Accelerometer Posture (CHAP) Method for Classifying Sitting Patterns from Hip Accelerometers: A Validation Study in Older Adults
Mikael Anne*, **Supun Nakandala***, Marta M. Jankowska, Dori Rosenberg, Fatima Tuz-Zahra, John Bellettiere, Jordan Carlson, Paul R. Hibbing, Jingjing Zou, Andrea Z. LaCroix, Arun Kumar, and Loki Natarajan
Medicine & Science in Sports & Exercise, 2021
(* Co-First Author)

Deep Learning Algorithms for Identifying Sedentary Behavior from Hip-worn Accelerometer Data
Supun Nakandala, Marta Jankowaska, Fatima Tuz-Zahra, John Bellettiere, Arun Kumar, and Loki Natarajan
Journal for the Measurement of Physical Behavior, 2021

Query Optimization for Faster Deep CNN Explanations
Supun Nakandala, Arun Kumar, and Yannis Papakonstantinou
SIGMOD Record 2020 **(SIGMOD Research Highlight Award)**

Incremental and Approximate Computations for Accelerating Deep CNN Inference
Supun Nakandala, Kabir Nagrecha, Arun Kumar, and Yannis Papakonstantinou
TODS 2020 **(Invited Paper)**

Workshop and Demo Publications

Compiling Classical ML Pipelines into Tensor Computations for One-size-fits-all Prediction Serving
Supun Nakandala, Gyeong-In Yu, Matteo Interlandi, and Markus Weimer
NeurIPS 2019 MLSys Workshop

Cerebro: Efficient and Reproducible Model Selection on Deep Learning Systems
Supun Nakandala, Yuhao Zhang, and Arun Kumar
SIGMOD 2019 DEEM Workshop

Demonstration of Krypton: Optimized CNN Inference for Occlusion-based Deep CNN Explanations
Allen Ordookhanians, Xin Li, **Supun Nakandala**, and Arun Kumar
VLDB 2019 Demo | MLSys 2019 Demo

Materialization Trade-offs for Feature Transfer from Deep CNNs for Multimodal Data Analytics
Supun Nakandala, Arun Kumar
MLSys 2018 Short paper

Anatomy of the SEAGrid Science Gateway
Supun Nakandala, Sudhakar Pamidigantam, Suresh Marru, Marlon Pierce
NSF XSEDE 2016

Pre-Prints

Nautilus: An Optimized System for Deep Learning-based Active Transfer Learning
Supun Nakandala and Arun Kumar
Under Preparation

Intermittent Human-in-the-loop Model Selection using Cerebro: A Demonstration
Liangde Li, **Supun Nakandala**, and Arun Kumar
Under Submission

Research Impact

Microsoft open-sourced Hummingbird system and uses it in ONNX ML Tools 2020
Ideas from project CEREBRO integrated into MADlib/Greenplum by VMWare 2019
CEREBRO system is being used by behavioral science researchers at UCSD 2019
"Gendered Conversation in a Social Game-Streaming Platform" paper gains lot of media attention and creates awareness about the bleak issue of sexism in online game streaming platforms 2017
APACHE AIRAVATA science gateways middleware and the SEAGRID science gateway are widely used by computational science researchers to execute and manage computational jobs on university clusters and national supercomputing infrastructure 2017

Patents

Pending US Patent Application: *Query Optimization for Deep Convolutional Neural Network Inferences*
Arun Kumar and Supun Nakandala

Pending US Patent Application: *Accelerating Inference of Traditional ML Pipelines*

with Neural Network Frameworks
Matteo Interlandi, Markus Weimer, Saeed Amizadeh, Konstantinos Karanasos, Supun Nakandala, Karla J. Saur, Carlo Aldo Curino and Gyeongin Yu

Service

Program Committee:
VLDB: 2022

External Reviewer:
VLDB: 2019

Mentoring Student Research Projects:
Liangde Li, MS UCSD 2021
Allen Ordookhanians, MS UCSD 2019
Xin Li, MS UCSD 2019

Advitya Gemawat, BS UCSD 2021
Kabir Nagrecha, BS UCSD 2021

Scholarships and Awards

Student grant to attend OSDI 2020	USENIX - 2020
NSF travel award to attend SIGMOD 2019	NSF - 2019
Gold Medal for the Best Academic Performance	University of Moratuwa - 2015
Travel award to attend South Asia Workshop on Research	NUS Singapore - 2014
Mahapola Higher Education Merit Scholarship	Govt. of Sri Lanka - 2010

Technical Talks

Cerebro: A Data System for Optimized Deep Learning Model Selection
VLDB 2020; UCSD CNS Research Review 2020; SIGMOD 2019

A Tensor Compiler for Unified Machine Learning Prediction Serving
OSDI 2020; Microsoft Gray Systems Lab 2019; Google Brain ML+Compiler Reading Group 2021 (Invited)

Vista: Optimized System for Declarative Feature Transfer from Deep CNNs at Scale
SIGMOD 2020; UCSD CNS Research Review 2018

Incremental and Approximate Inference for Faster Occlusion-based Deep CNN Explanations
SIGMOD 2019

Gendered Conversation in a Social Game Streaming Platform
AAAI ICWSM 2017; Indiana University Center for Complex Network and Systems Research 2017 (Invited)

Apache Airavata Sharing Service: A Tool for Enabling User Collaboration in Science Gateways
PEARC 2017

Apache Airavata Security Manager: Authentication and Authorization Implementations for a Multi-tenant e-Science Framework
IEEE e-Science 2016

Anatomy of SEAGrid Science Gateway
PEARC 2016