Aditya Krishnan

+1 412-623-9519 • aditya.krishnan94@gmail.com • in arkrishn • Google Scholar

Research Interests

Similarity search, retrieval augmented generation, scalable machine learning, information retrieval.

Professional Experience

Microsoft **Mar 2025 – Present**
Remote *Senior Software Engineer*

- Working on scaling core vector indexing algorithms for M365 Copilot search experiences within Substrate.

Pinecone Systems **Jan 2024 – Oct 2024**
New York City *Senior Research Scientist*

- Developed a novel query-routing mechanism for IVF-style vector indexes, reducing scanned data by 25% vs. Google ScaNN algorithm (published in NeurIPS 2025).
- Implemented a new negative sampling method for text cross-encoders, improving document retrieval accuracy and surpassing industry rerankers (e.g., Cohere-Rerank-V3).

Oct 2022 – Jan 2024
Research Scientist

- Designed Pinecone's winning solution for the NeurIPS 2023 Big ANN Challenge (Out-of-Distribution track) in Rust, achieving 35K+ QPS on a 10M-sized index using 8 vCPUs and 16 GB RAM.
- Led research on vector quantization for Pinecone Serverless using SIMD-optimized Rust kernels, achieving a 2× improvement in cached data utilization.

Pinecone Systems **May 2021 – Aug 2021**
Remote *Science Intern*

- Researched quantization techniques for similarity search advied by Edo Liberty (CEO and Founder).

Education

Johns Hopkins University, Whiting School of Engineering **Sep 2018 – Sep 2022**
Baltimore, MD *Ph.D. in Computer Science*
Advisor: Vladimir Braverman
Thesis: *Fast and Memory-Efficient Algorithms for Matrix Spectrum Approximation*

Carnegie Mellon University, School of Computer Science **May 2017 – May 2018**
Pittsburgh, PA *M.S. in Computer Science*
Advisor: Anupam Gupta
Thesis: *Pricing Online Metric Matching Algorithms on Trees*

Carnegie Mellon University, School of Computer Science **Aug 2013 – May 2017**
Pittsburgh, PA *B.S. in Computer Science (Minor: Engineering Studies)*

Honors and Awards

2022: JHU MINDS TRIPODS Data Science Fellowship (awarded to ∼5 students across two schools per cycle).

2022: NeurIPS Top Reviewer (less than 10% of reviewers).

2018: JHU Computer Science Department Fellowship (awarded to 2 students in an incoming class of 50+).

Technical Skills

Languages: Rust; Python.

Libraries: PyTorch; Distributed Data Parallel; FSDP; SciKit-learn.

Publications

Authors appear in alphabetical order. * denotes equal contribution.

NeurIPS 2025: Optimistic Query Routing for Maximum Inner-Product Search. With Sebastian Bruch and Franco Maria Nardini.

ICALP 2023: Lower Bounds for Pseudo-Deterministic Counting in a Stream. With Vladimir Braverman, Robert Krauthgamer, and Shay Sapir.

STOC 2022: Sublinear Time Spectral Density Estimation. With Vladimir Braverman and Christopher Musco.

ACML 2021: Lifelong Learning with Sketched Structural Regularization. With Haoran Li, Jingfeng Wu*, Soheil Kolouri*, Praveen K. Pilly, and Vladimir Braverman.

COLT 2021: Near-Optimal Entrywise Sampling of Numerically Sparse Matrices. With Vladimir Braverman, Robert Krauthgamer, and Shay Sapir.

ICML 2020: Schatten Norms in Matrix Streams: Hello Sparsity, Goodbye Dimension. With Vladimir Braverman, Robert Krauthgamer, and Roi Sinoff.

WINE 2020: Competitively Pricing Parking in a Tree. With Max Bender, Jacob Gilbert, and Kirk Pruhs.

APPROX 2018: On Sketching the q to p Norms. With Sidhanth Mohanty and David P. Woodruff.

Talks

2022: Sublinear Time Spectral Density Estimation, STOC, Rome, Italy.

2018: Sublinear Time Spectral Density Estimation, JHU CS Theory Seminar, Baltimore.

2020: Schatten Norms in Matrix Streams: The Role of Sparsity, ICML.

2019: Schatten Norms in Matrix Streams: The Role of Sparsity, JHU CS Theory Seminar, Baltimore.

2018: Pricing Online Metric Matching Algorithms on Trees, CMU Theory Seminar, Pittsburgh.

Academic Service

Reviewer: NeurIPS 2021–2024; ICML 2021–2024; ICLR 2022–2024; STOC 2021–2022; SODA 2021; PODS 2020.

Seminar Organizer: JHU Theory Seminar (2021, 2022).

Teaching Assistant: Introduction to Algorithms (Fall 2019, Spring 2020, Spring 2022); Approximation Algorithms (Spring 2021).

References

Edo Liberty, CEO and Founder, Pinecone — edo@edoliberty.com

Christopher Musco, Assistant Professor, New York University — cmusco@nyu.edu

Vladimir Braverman, Professor, Rice University — vb21@rice.edu