

Arnav Sareen

919-888-3818 | asareen2@charlotte.edu | [linkedin.com/in/arnavsareen](https://www.linkedin.com/in/arnavsareen) | github.com/aesareen | aesareen.github.io/

EDUCATION

University of North Carolina at Charlotte

Charlotte, NC

Bachelor of Science in Computer Science (AI Concentration), Bachelor of Science in Data Science

2023 – 2027

Honors: Levine Scholar (*Premier Merit Scholarship*), Chancellor's List — Fall 2023 - Spring 2026

GPA: 4.0/4.0

Tampere University

Tampere, FI

Exchange Studies

Jan. 2026 – May 2026

RESEARCH EXPERIENCE

Fulbright Canada MITACS Globalink Research Intern

May 2026 –

SHARE Lab — Advisor: Dr. Sharon Ferguson

Waterloo, ON, CA

- Investigating factors influencing ML/AI career persistence by scraping and analyzing university curricula, mapping experiential-learning gaps, and modeling how academic pathways shape students' perceived career fit.

AI & HPC Researcher

August 2023 – Present

Data Intelligence Research (DIR) Lab — Advisors: Dr. Dong Dai, Chris Egersdoerfer

Charlotte, NC

- Evaluated LLMs for interpreting complex HPC I/O logs using a scientific-paper-grounded RAG system, reducing hallucinations in performance-diagnosis responses.
- Improved LLM consistency and diagnostic accuracy for I/O trace-log analysis by 80% through prompt optimization, structured evaluation, and iterative benchmarking.
- Built an LLM evaluation and benchmarking interface to track RAG pipeline changes, enabling systematic comparison of retrieval, prompting, and generation strategies.

MLSys Research Assistant

June 2022 – August 2022

PICTure Research Group — Advisors: Dr. Xipeng Shen, Dr. Yuanchao Xu

Raleigh, NC

- Developed an optimized retraining strategy for learned indexes under high-write workloads, reducing workload-induced accuracy degradation by 60%.
- Presented findings on learned-index retraining efficiency at the NCSSM Mentorship Research Symposium.

PUBLICATIONS

C. Egersdoerfer, **A. Sareen**, J. L. Bez, S. Byna, D. D. Xu and D. Dai, IOAgent: Democratizing Trustworthy HPC I/O Performance Diagnosis Capability via LLMs, *2025 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, Milano, Italy, 2025, pp. 322-334, doi: 10.1109/IPDPS64566.2025.00036

C. Egersdoerfer, **A. Sareen**, J. L. Bez, S. Byna, and D. Dai. ION: Navigating the HPC I/O Optimization Journey using Large Language Models, in *Proceedings of the 16th ACM Workshop on Hot Topics in Storage and File Systems (HotStorage)*, Santa Clara, CA, USA, 2024, pp. 86–92. <https://dl.acm.org/doi/10.1145/3655038.3665950>

SELECTED AWARDS AND FELLOWSHIPS

Barry M. Goldwater Scholarship: Named a 2026 Goldwater Scholar—the preeminent national honor for STEM researchers—for demonstrating exceptional potential in artificial intelligence and parallel computing. *Awarded 2026*

George Barthalmus Undergraduate Research Award: Research Award developed to promote early involvement in the research process through support of sophomores in a research project of their design *Awarded 2024*

Levine Scholarship: The Levine Scholars Program is UNC Charlotte's premier fully-funded undergraduate merit scholarship, developing intellectual curiosity, commitment to community service and ethical leadership. *Awarded 2023*

National Cyber Scholar with Honors: Competed for a top score against thousands of students across the nation by solving cybersecurity problems. Additionally, completed and passed the GIAC GFACT exam and obtained one of the highest 500 scores within the National Cyber Scholar cohort. *Awarded 2023*

TEACHING / MENTORING

Code In Place Section Leader

April 2026 – May 2026

Stanford School of Engineering

Remote

- Selected as one of 800 section leaders from 2,000+ applicants; taught weekly Python sections for 10+ beginner programmers and provided individualized debugging support.

Teaching Assistant

August 2025 – December 2025

UNC Charlotte College of Computing & Informatics

Charlotte, NC

- Developed and graded hands-on assignments on clustering, classification, and regression for a 70+ student Data Mining course.
- Led office hours and code reviews, helping students debug implementations of classical machine-learning algorithms and strengthen applied modeling skills.

E3 Technical Fellow

July 2025 – December 2025

Codepath

Charlotte, NC

- Mentored 20+ students in data structures, algorithms, and technical interview preparation through individualized coaching sessions.
- Conducted mock technical interviews and provided targeted feedback on algorithmic reasoning, code clarity, debugging, and communication.

INDUSTRY EXPERIENCE

Open Source Contributor

October 2025 – Present

Codepath

Remote

- Contributing to large-scale open-source codebases, including Apache Superset and GitLab, collaborating with maintainers through pull requests and code reviews.

Data Science and Analytics Intern

May 2025 – December 2025

Trane Technologies

Davidson, NC

- Developed a hybrid recommendation engine combining matrix factorization and deep-learning collaborative filtering to support customer/product recommendation workflows.
- Designed a RAG system using AWS Bedrock Agents, Lambda, and Redshift to generate contextualized responses from enterprise data.

Data Science and AI Engineer Intern

May 2024 – September 2024

Charlotte Works

Charlotte, NC

- Developed a custom AI assistant and chatbot through the integration of advanced RAG techniques such as HyDE and Hybrid Retrieval.
- Analyzed labor market data through Python and statistical tools to uncover key labor trends and growth opportunities in Mecklenburg County.

SERVICE & LEADERSHIP

Fellowship Coordinator

June 2025 – Present

Strength in Numbers

Charlotte, NC

- Created Charlotte's first sports analytics program for high school students, providing underrepresented learners with technical training, industry exposure, and college application guidance.

Editor-in-Chief & Web Developer and Media Editor

June 2024 – Present

ETHEL Undergraduate Research Journal

Charlotte, NC

- Re-establishing the Charlotte's premier journal for undergraduate research, spearheading administrative duties, student outreach efforts, and the digital version of the journal.

TECHNICAL SKILLS AND OTHER AFFILIATIONS

Courses: Data Structures & Algorithms, Computer Systems, Database Design, Operating Systems & Networks, Natural Language Processing, Artificial Intelligence, Machine Learning, Data Mining, Applied Regression, Multivariate Analysis, Multivariable Calculus, Computer Architecture, Fine-tuning Large Language Models, Complex Networks

Languages: Python, Java, SQL, C/C++, SAS, JavaScript, HTML/CSS, RISC-V

Developer Tools: Git, VS Code, AWS Sagemaker, AWS Bedrock, AWS Redshift, BigQuery, PyCharm, IntelliJ

Libraries + Tools: Pandas, Polars, NumPy, Matplotlib, DsPy, ChromaDB, Plotly, Scikit-Learn, spaCy, BeautifulSoup, Streamlit, Selenium, Flask, Git

Clubs and Organizations: Sports Analytics Club Member, Charlotte AI Research Club Member