Daniel Öman

dsoman24@gmail.com | (470) 553-5299 | Atlanta, GA | github.com/dsoman24 | linkedin.com/in/daniel-s-oman

Education

Georgia Institute of Technology — ***August 2021 – May 2025 (expected)***

B.S. Computer Science, concentrations in Intelligence (AI/ML) and Theory — 3.96/4.0 GPA

Relevant Coursework: Data Structures & Algorithms, Machine Learning, Deep Learning, Graduate Algorithms, Computer Organization & Programming, Probability & Statistics, Automata & Complexity, Number Theory

Experience

Google — Kirkland, WA

Software Engineering Intern — ***May – August 2024***

- Design and implement a distributed load testing framework using C++ and Python to gauge scalability of a novel data streaming service and write-ahead log within Google BigQuery's core infrastructure.
- Implement continuous runs of the framework as a custom workflow to be used in development to detect bottlenecks and prevent latency and throughput regressions before reaching production, leading to a projected 20% reduction in server request latency.
- New load sampling framework reduces test implementation time by 50%, enabling faster iteration cycles for performance testing.
- Developed custom monitoring dashboards, allowing for 4x faster detection of critical performance regressions.

Georgia Tech Efficient and Intelligent Computing Lab — Atlanta, GA

Undergraduate Research Assistant — ***January – May 2024***

- Contributed to a PyTorch toolkit to train distributed Graph Neural Networks (GNNs) for applications with multiple disjoint large graph datasets, such as electronic design automation analysis.
- Built a user-friendly modular data loading and transfer API and implemented the GraphSAGE GNN forward propagation and graph vertex embedding algorithm, improving model accuracies by an average of 15%.

Georgia Tech College of Computing — Atlanta, GA

Undergraduate Teaching Assistant (Homework Lead) — ***August 2022 – May 2024***

- Managed a team of 40 TAs in the development and grading of 12 homework assignments for over 800 students per semester as TA Homework Lead for CS 1331: Intro to Object-Oriented Programming (Java) under Prof. Richard Landry and Dr. Aibek Musaev.
- Led weekly recitations for 50 students and helped students with problem-solving and debugging during one-on-one office hours.

Google — Sunnyvale, CA

Software Engineering (STEP) Intern — ***May – August 2023***

- Implemented and tested an efficient parallel-processing data pipeline being used in production to train machine learning models that predict Google Workspace account upgrade, downgrade, and churn behaviors.
- Built pipeline using FlumeJava, a Java MapReduce framework, to extract and aggregate 70+ web domain level ML features from a database containing the HTML of more than 500 billion web pages, increasing customer coverage in the feature store by 20%.
- Engineered a scalable and extensible data aggregation architecture by applying advanced OOP design patterns that reduced feature implementation time by over 50% and provided an intuitive interface for future feature store contributions.
- Refactored pipeline to improve reliability by developing a system to flush intermediate output to a Spanner database across 100k+ threads during a full table scan, preventing data loss by storing more than 7 days of data progress during each pipeline execution.

Georgia Tech Financial Services and Innovation Lab — Atlanta, GA

Undergraduate Research Assistant (Team Lead) — ***January – May 2023***

- Led a team of 4 researchers developing a Python pipeline to perform sentiment analysis on earnings calls transcripts on 12 electric vehicle companies using the large language model FinBERT and natural language processing library NLTK.
- Developed a custom web scraper using Beautiful Soup to extract over 70 earnings call transcripts from The Motley Fool.

Projects

Hemodynamics Calculator | *JavaScript, ReactJS, MongoDB, Express, NodeJS* — ***August 2023 – April 2024***

- Developed the Hemodynamics Calculator, a full-stack MERN application for the Emory University School of Medicine used by over 10 clinicians to reduce blood flow measurement error daily, critically impacting more than 1,000 cardiac ICU patients a year.
- Placed 3rd out of 50 teams in the Georgia Tech CS Capstone Expo, presenting to 40+ industry professionals and professors.

Machine Learning Soccer Prediction | *Python, sklearn, PyTorch, NumPy, Matplotlib, Seaborn* — ***August – December 2023***

- Worked on a team of 5 to build and train logistic regression, random forest, and artificial neural network models using Scikit-Learn and PyTorch to predict soccer match outcomes with 70% accuracy, beating benchmark betting odds data by 8%.
- Built feature engineering strategies and conducted hyperparameter tuning to reduce overfitting, improving accuracy by ∼10%.

Skills

Programming Languages: Java, C/C++, Python, SQL, JavaScript, LaTeX

Frameworks: FlumeJava (MapReduce), JUnit, NumPy, Pandas, Scikit-Learn, PyTorch, ReactJS, NodeJS, Flask

Tools: Git, Mercurial, Bazel, Protobuf, gRPC, Spanner, MySQL, MongoDB