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 Courses: Data Structures & Algorithms, Machine Learning, Deep Learning, Computer Organization & Programming
Activities: Undergraduate Teaching Assistant (Lead TA), Delta Chi Fraternity (Secretary, Professional Development Chair)

Experience

Ditch Atlanta, GA
Software Engineering Intern ***August 2024 – Present***

- Early-stage FinTech startup building a platform for automated debt repayment.
- Full-stack development using Google Cloud Run Functions in Python for the backend and Flutter with Dart for the frontend.

Google Kirkland, WA
Software Engineering Intern ***May 2024 – August 2024***

- Designed and implemented a custom distributed load testing framework using C++ and Python to benchmark the scalability of the streaming metadata change-log service within Google BigQuery's core storage infrastructure.
- Built continuous test runs as a development workflow, leading to 70% reduction in regressions before reaching production.
- Developed load sampling architecture to simulate production traffic by sending 10k+ requests per second to the read and write RPC endpoints on a variable number of BigQuery tables, exposing production bottlenecks.
- Designed a multi-threaded metric sampling system with C++ to compute, aggregate, and analyze latency, throughput, and error rate metrics across 25+ load sampler instances concurrently over multiple machines, yielding 50% increased benchmark accuracy.
- Led efforts to fix a critical service-level objective bug affecting BigQuery's storage metadata server by implementing request retry logic, eliminating the number of error spikes by ~90%.

Georgia Tech Efficient and Intelligent Computing Lab Atlanta, GA
Undergraduate Research Assistant ***January – May 2024***

- Contributed to a PyTorch toolkit used by 5+ Georgia Tech labs to train distributed Graph Neural Networks (GNNs) for applications with multiple large disjoint graphs, such as electronic design analysis and molecular modeling.
- 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%.

Google Sunnyvale, CA
Software Engineering (STEP) Intern ***May – August 2023***

- Designed, implemented, and tested an efficient parallel-processing data pipeline being used in production to provide features 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 10k+ processes during a full table scan, preventing up to 7 days worth of lost data for each pipeline run.

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, Scikit-Learn, 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, Express, NodeJS, Flask
Tools: Git, Mercurial, Bazel, Protobuf, gRPC, Spanner, MySQL, MongoDB