

# SOHAM DANDAPATH

[soham.dandapath@gmail.com](mailto:soham.dandapath@gmail.com) • +1 646-704-8790 • [linkedin.com/in/soham-dandapath](https://www.linkedin.com/in/soham-dandapath) • [github.com/12dash](https://github.com/12dash)

---

Applied AI Engineer delivering production GenAI and forecasting systems on agentic platforms. Drive RAG/LLM and time-series solutions from technical discovery to deployment, with \$7M+ in annual delivered impact and \$0.8B+ in active pipeline.

## PROFESSIONAL EXPERIENCE

---

**C3.ai** • Redwood City, CA *Jan 2024 to Present*

*Applied AI Engineer (Data Scientist)*

- Lead enterprise forecasting program for a major semiconductor customer with **\$0.8B estimated annual impact**, owning technical discovery, success-metric alignment, and end-to-end production deployment on the C3 agentic platform.
- Delivered **\$5M+ annual value** via yield forecasting for the world's largest berry producer and **\$2.3M annual value** via demand forecasting for the largest CPG company in Guatemala; owned modeling, validation, and rollout end-to-end.
- Led end-to-end build and deployment of a **RAG-based LLM system** for unstructured document retrieval, enabling low-latency, policy-compliant generation across enterprise sales and product documentation.
- Architected and own **C3 DRIPP**, the internal Python and CI/CD toolchain that cut deployment latency from hours to minutes, and **MetaML**, the orchestration platform for enterprise time-series deployments adopted across customer engagements.
- Direct release management for core forecasting packages, enforce software engineering and CI/CD standards, and **mentor data scientists** across customer-facing and internal applied ML initiatives.

**C3.ai** • Redwood City, CA *Jun 2023 to Aug 2023*

*Data Science Intern*

- Shipped an out-of-the-box hierarchical forecasting and reconciliation system on the C3 AI platform (post-hoc MinT/ERM and intrinsic DeepVAR-Hierarchical) ensuring cross-level forecast coherence.
- Integrated probabilistic forecasting with **Integrated Gradients-based explainability** to deliver uncertainty-aware, interpretable predictions within the platform.

**Charles and Keith** • Singapore *Jan 2022 to Jul 2022*

*Data Scientist*

- Built an image similarity engine with **95%+ accuracy** that improved global product matching and inventory decisions.
- Developed tree-based regression models for seasonal demand forecasting and strategic planning.

**Shopee** • Singapore *May 2021 to Jul 2021*

*Data Engineering Intern*

- Built HDFS compression tooling that reduced storage usage by **90%** and managed daily ingestion pipelines orchestrated through Airflow.

**Seagate Technology** • Singapore *Aug 2020 to Dec 2020*

*Machine Learning Engineer Intern*

- Built neural and tree-based models forecasting custom hard-drive test times; integrated into a factory web app for live pipeline visibility and accurate delivery commitments.

## RESEARCH PROJECTS

---

**Controllable Face Modeling with Causal DAGs and BiGAN:** Combined DAG-guided causal representation learning with BiGANs to enable attribute-level latent interventions while preserving identity.

**DNN-FES and FOLE (Undergraduate Thesis):** Designed DNN-FES to extract human-readable fuzzy rule sets from deep networks; introduced FOLE, a consistency loss preserving logical coherence to improve interpretability and fidelity.

## EDUCATION

---

**Columbia University** • M.S. Computer Science, ML Concentration • GPA 4.06 / 4.00 *Sep 2022 to Dec 2023*

Coursework: ML, Deep Learning, Reinforcement Learning, Causal Inference I and II, Learning Theory, Databases.

**Nanyang Technological University** • B.Eng. Computer Science (Math Minor) • GPA 4.66 / 5.00 *Aug 2018 to May 2022*

First Class Honors; Dean's List 2021 to 2022; Research Assistant at Nanyang Business School under Prof. Moksh Matta.

## SKILLS

---

**GenAI and ML:** RAG, LLMs, Agentic Workflows, Semantic Retrieval, Probabilistic Time-Series Forecasting, Causal Inference, Distributed Training

**Engineering:** Python, PyTorch, Spark, AWS, GCP, CI/CD, Release Management, Software Architecture