

Hassan Dawy

hdawy@bu.edu | 617-530-0573 | Boston, MA | [linkedin.com/in/hassandawy](https://www.linkedin.com/in/hassandawy) | github.com/HassanDawy

EDUCATION

Boston University, Faculty of Computing and Data Sciences (CDS)

Expected Dec 2027

BS in Data Science

GPA: 3.90/4.00

Relevant coursework: Theory and Applications of LLMs (graduate), Data Engineering at Scale (graduate), Data Structures and Algorithms, Machine Learning and AI, Linear Algebra, Probability and Statistics

WORK EXPERIENCE

Machine Learning Engineer Intern - noon (UAE)

Jun 2026 - Aug 2026

- Built a seller revenue forecasting system from scratch, producing 12-month quantile forecasts for **28,000+** sellers and powering financing decisions worth **~\$200M**
- Designed an out-of-time backtest over 8 quarters (**890K+** forecasts), cutting WAPE **15%** vs. the best baseline
- Engineered a leakage-safe 4.9M-row, 46-feature training table from BigQuery with parameterized SQL
- Deployed it as a self-serve FastAPI service with event-driven batch workers, Postgres caching, and access control

Machine Learning Engineer & Project Manager - Benevolent Bandwidth Foundation

Feb 2026 - Present

- Leading 6 people on a breast cancer screening tool, using MobileNetV2 (2.3M params) for on-device inference
- Run containerized Vertex AI training on NVIDIA T4 GPUs with tf.data, GCS checkpointing, and pytest CI
- Fixed a BatchNorm failure mode; benchmarked input sizes (288px: +0.012 AUROC for 15% more train time)
- Reached **0.83** AUROC and 82% sensitivity on 75K+ mammograms with patient-grouped splits to prevent leakage

Full Stack Developer - Boston University CDS Clinic

Jan 2026 - May 2026

- Built a real-time ETL pipeline that uses cron jobs to load minute-level AWS S3 reports into normalized SQLite schemas, served to the app through a FastAPI backend
- Built a React/TypeScript PWA on Google Maps showing live garage availability for **2,500+** potential users
- Ran bi-weekly demos with BU IS&T stakeholders to align on data schemas, data contracts, and feature priorities

PROJECTS

av-agents: Autonomous Multi-Agent Coding Pipeline | TypeScript, LLM APIs, GitHub Actions Aug 2026 - Sep 2026

- Built a multi-agent system that wrote most of Anchor Vision, an ~80k-line TypeScript app alerting day traders when their charts drift from their written strategy; deployed its backend on AWS (ECS Fargate, Aurora, SQS, Cognito)
- Designed a state-machine orchestrator (planner, tool-calling coder, build gate, two cross-vendor reviewers, merge on double approval) with human escalation; merged **57** tasks in 13 days
- Instrumented token and cost tracing for all **592** LLM calls; profiling showed the coder drove 76% of spend because its tool loop re-sent full context each step, so I built cost-gated context compaction
- Added accountability controls: an infrastructure-vs-task failure classifier, dry-run push preflight, and per-task, per-tick, and 24-hour spend caps backed by a per-call ledger
- Fixed a non-converging review loop (~52% rejection rate) with issue memory, persistent verdicts, and a stopping rule

NYT Connections Solver | Python, PyTorch, DeBERTa-v3

Apr 2026

- Trained a DeBERTa-v3 group classifier with beam search over 1,820 candidates; traced a misleading 75% accuracy to an all-negative model
- Ablated sampling ratio, unfreezing, learning rate, batch size, and beam width; the scorer was the bottleneck
- Reached **9.9%** puzzle wins and **16.2%** group accuracy on 131 held-out puzzles vs. 0% for both baselines

Facebook Graph Analysis | Rust

Nov 2024

- Implemented graph algorithms in Rust on real Facebook ego-network data: BFS for average path distances, closeness centrality to find influential nodes, and Jaccard similarity over mutual friends to find communities

SKILLS

- **Languages:** Python, TypeScript, SQL, Rust, Java
- **ML:** PyTorch, TensorFlow/Keras, scikit-learn, model evaluation and ablations, LLM agents and tool-calling
- **Infra:** GCP (Vertex AI, GCS, BigQuery), AWS, Docker, GitHub Actions, FastAPI, PostgreSQL, pytest