Sami Hatoum

07741 157779 — shatoum101@gmail.com — linkedin.com/in/sami-hatoum — github.com/Sami-Hat — www.samihatoum.dev

Education

BSc Computer Science, University of St Andrews *First Class Honours, Dean's List, 2021 - 2025*
MSc Artificial Intelligence and Medical Imaging, University College London *Predicted First Class, 2025 - 2026*
AWS Certified Cloud Practitioner, Amazon Web Services *2023*

Experience

PricewaterhouseCoopers (PwC) *Jun 2023 – Jul 2023*
Software Consulting Intern (AI)

- Built a client sentiment analysis prototype using a **PyTorch-based BERT** model, achieving a **91% F1 score** on an imbalanced customer review dataset (3:1 negative to positive)
- Containerised the model with Flask API wrapper and deployed to **AWS EC2** with auto-scaling, load tested to handle **150 req/s** with average **400ms response time**
- Conducted monitoring via CloudWatch and established a model retraining pipeline triggered by accuracy degradation below **85% threshold**
- Optimised preprocessing pipeline with pandas vectorisation, reducing daily batch **processing time by 60%**

Projects

Database Query Optimisation Tool

- A **PostgreSQL performance analyser** that parses **query ASTs** and evaluates **EXPLAIN plans** to recommend indexes, reducing query execution time from **2.4s to 420ms** on table scans over **500K rows**
- Analysed **200+ production queries** from a **five-million-row database** across 12 tables, identifying **17 missing indexes** that reduced average query time by **65%** and eliminated full table scans on **89% of frequently-run queries**
- Deployed on **AWS** using **ECS** with **RDS PostgreSQL**, implementing **CloudWatch monitoring** with alerts for query latency over 5s and configuring **auto-scaling groups** to handle traffic spikes during batch analysis runs
- Built interactive **D3.js dashboard** visualizing query execution plans as **flame graphs** and **index coverage heatmaps**
- Implemented **cost-based recommendation engine** that prioritises indexes by expected performance gain, preventing over-indexing that degrades write performance by estimating **15-20% INSERT/UPDATE overhead** per additional index

Food Waste Managemer

- A **React Native** food tracking app to help reduce household food waste, with 30+ beta users testing over 3 months
- Deployed Node.js REST API on **Vercel** with **RDS** (PostgreSQL) backend with automated daily backups
- Integrated **Recipe APIs** with **Redis caching**, reducing average response time from **1.8s to 120ms**
- Established push notifications using **AWS SNS** triggering on approaching expiration dates, achieving **68% user engagement rate** with 2.3 average notifications per user per week
- Conducted UI/UX accessibility review using **Android Accessibility Scanner** and verified consistent 60 FPS rendering on mid-tier devices using **Android Studio Profiler**

Deep Learning Medical Image Segmentation

- A **PyTorch medical image segmentation system** for prostate anatomy in T2-weighted MRI scans using a U-Net with attention and residual blocks.
- Designed a data augmentation pipeline with geometric and intensity transforms to address class imbalance
- Benchmarked against **nnU-Net** and **MONAI** using Dice, IoU, and Hausdorff distance.

Cryptographically Secure File Server

- An end-to-end encrypted **Java** file server with **zero-knowledge** design using AES 256, RSA 2048 key exchange, and HMAC SHA256 integrity verification
- Designed **capability based access control** that hides all plaintext identifiers and content from the server
- Added **secure multi party computation** for encrypted file sharing with a **zero-trust** server model
- Load tested with **JMeter** at 50 concurrent sessions and 200 operations/second, reaching **180 ms retrieval latency** and **450 ms upload latency** for files up to 10 MB
- Assessed security against **OWASP Top 10** with penetration testing for man in the middle and replay attacks

Skills

Core: Java, JavaScript, TypeScript, Python, SQL
Frameworks: React, Next.js, Node.js, Flask, React Native
Cloud & DevOps: AWS, Docker, Git, Kubernetes, Jira, CI/CD (GitHub Actions, Jenkins)
AI & Data: TensorFlow, PyTorch, Pandas, Scikit-learn