Sami Hatoum

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

Education

University of St Andrews 2021 – 2025
BSc Computer Science *First Class Honours, Dean's List*

University College London 2025 – 2026
MSc Artificial Intelligence and Medical Imaging *Predicted Distinction*

Experience

PricewaterhouseCoopers (PwC) *Jun 2024 – Aug 2024*
Software Consulting Intern (AI)

- Built a **sentiment analysis model** using **PyTorch BERT** to automate customer feedback classification for a retail client. Achieved **91% F1 score** on an imbalanced dataset (3:1 negative to positive), processing **10,000 reviews daily**
- Engineered containerised deployment with **Flask API** on **AWS EC2** with auto-scaling across multiple instances. Collaborated with 3 senior consultants to validate **sub-500ms latency** under load testing
- Set up **CloudWatch** monitoring dashboards and wrote a retraining script triggered when weekly evaluation accuracy dropped below **85%**
- Optimised the preprocessing pipeline with **pandas vectorisation**, reducing daily batch processing time by **60%**

PricewaterhouseCoopers (PwC) *Aug 2023 – Sep 2023*
Software Engineering Intern

- Developed an internal **client onboarding dashboard** in **React** and **Node.js** that replaced a manual spreadsheet workflow for a financial services team, reducing onboarding processing time from **3 days to 4 hours**
- Designed and implemented a **PostgreSQL** data model with role-based access control, integrating with PwC's internal **SSO authentication** system
- Wrote **CI/CD pipelines** in **GitHub Actions** with automated testing and staged deployments to **AWS**, achieving **90%+ code coverage** across 45 unit and integration tests

Projects

Database Query Optimiser

- Developed a **PostgreSQL analysis tool** that parses **query ASTs**, evaluates **EXPLAIN plans**, and recommends indexes. Reduced query execution time from **2.4s to 420ms** on table scans over **500K rows**
- Reviewed **200+ production queries** across a **five-million-row database** (12 tables), identifying **21 missing indexes** that reduced average query time by **65%**
- Deployed on **AWS ECS** with **RDS PostgreSQL**, implementing **CloudWatch** alerts and **auto-scaling groups** for batch analysis traffic spikes

Selective Prediction for Prostate Cancer Segmentation (MSc Dissertation)

- Trained **nnU-Net** (5-fold cross-validation, 9-class segmentation) on multi-institutional prostate MRI, achieving **0.82 mean foreground DSC** with per-organ scores exceeding **0.90** on major structures
- Implemented **Monte Carlo Dropout** inference (T=20 forward passes) with entropy, mutual information, and coefficient of variation aggregation to produce per-scan clinical deferral decisions
- Built a **selective prediction framework** that maintained **0.89 DSC at 85% coverage**, automatically flagging unreliable cases for clinical review and reducing potential silent failures in treatment planning

Hierarchically Weighted Image Segmentation (Group Project)

- Designed a custom **hierarchical loss function** encoding radiosensitivity scores, anatomical proximity to isocentre, and asymmetric penalty weighting for prostate **IMRT** planning
- Trained a **3D Residual Encoder U-Net** on **589 multi-institutional T2-weighted MRI** volumes (7 institutions, 1.5T and 3T scanners) achieving **0.886 mean foreground DSC**; improved **NVB segmentation from 0.630 to 0.659** DSC over unweighted baseline
- Conducted **ablation studies** across 3 loss functions and 3 weighting schemes, isolating contributions of sensitivity and distance error weighting

Skills

Programming Languages: Python, Java, JavaScript, TypeScript, SQL
Frameworks: React, Next.js, Node.js, Flask, React Native
Cloud & DevOps: AWS (Certified Cloud Practitioner), Docker, Git, Kubernetes, Jira, CI/CD
AI & Machine Learning: TensorFlow, PyTorch, BERT, NLP, U-Net, Scikit-learn, Pandas, NumPy