Hisham Barakat

AI Engineer — Infra & Generative AI — Built & Scaled GenAI Platform from 3 → 30 Models

Hishambarakat16@gmail.com — +86 166 0152 0270 — +973 3302 5065 — LinkedIn — Portfolio

Education

Master in Computer Science, Peking University   Sept 2023 – Present
- **Courses:** Computational Vision, Reinforcement Learning, Computer Networking, Advanced Systems.
- Optimize Jittor backend for TensorLayerX's multi-backend AI framework.

Bachelor of ICT, Bahrain Polytechnic   Sept 2017 – May 2021

Technical Skills

Languages: Python, C++, Java, SQL, YAML

Technologies: PyTorch, TensorRT, Diffusion, NLP, CV, Hugging-Face, Transformers, LangChain, PEFT, BitsAndBytes, OpenAI, Anthropic/Claude, FastAPI, Uvicorn, Flask, Docker, Kubernetes, AWS, GCP, Power-BI, Git, CUDA

Experience

Lead AI Engineer (collaborative project), Jiangsu Huaxiang Digital Technology Group Co.,Ltd   Sept 2024 – Present
- **Spearhead** a scalable image & video generation platform; grow model count **3 → 30** and serve them via **FastAPI/Uvicorn REST APIs**.
- **Boost** GPU-bound **image-generation** throughput by **3.5×** and cut average latency **60%** while maintaining **100%** success, leveraging per-model **CUDA streams**, dynamic batching, and **CPU/IO multithreading**.
- **Fine-tune Whisper** (Arabic ASR), **Spark-TTS** (Arabic TTS), and **Qwen-3** LLM with **LangChain**, **Transformers**, **LoRA/PEFT**, and **Unsloth**.
- **Deploy** containerized APIs on cloud with **Docker**; add auto-scaling, OSS-stack logging, and real-time dashboards in **Power BI**.
- **Mentor 3 junior engineers**, drive code reviews and best-practice guides, and lift team velocity by **20%**.

IP Engineer, Huawei Technologies Bahrain W.L.L   Sept 2022 – Aug 2023
- Developed **Python-driven automation** scripts that cut network deployment time by over **50%**.
- Performed network **patching**, **deployment**, and maintenance for multiple clients under strict SLAs, ensuring **100%** network availability.
- Configured and optimized routing protocols including **IPv6**, **BGP**, **EIGRP**, **OSPF**, and **RIP**.

Data Science & AI Engineer, Tatweer Petroleum   Sept 2021 – Aug 2022
- Collaborated with two production teams to aggregate and clean over **1M data points** from **6** systems using AWS.
- Conducted diagnostic analyses to uncover **root causes of oil loss**, driving process changes that saved over $**100K**.
- Developed and automated **15+ Power-BI reports**, reducing manual extraction workload by **10%**.
- Presented findings and recommendations to stakeholders, fostering data-driven decision-making.

Solutions Architect Intern, Amazon Web Services   Feb 2021 – Jun 2021
- Implemented a cost-effective, end-to-end serverless face-mask detection dashboard using **AWS Rekognition**, **Lambda**, **S3**, and **DynamoDB**—cutting expenses by **95%**.
- Deployed **Amazon DeepLens** for facemask detection with over **90%** accuracy and integrated **Lambda** functions with other AWS services.
- Led a team of **5** developers, communicated progress to customers, and monitored system health via **CloudWatch**.

Awards and Certifications

- **AWS Certified Machine Learning - Specialty**
- **Commendation Letter from Huawei Bahrain CEO**
- **Huawei ICT Competition 2019:** 3rd Place Global; 2nd Place Middle East & National