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
- Architect a scalable image & video generation platform, expanding from **3 to 30 models** for local use and **API based** deployment, by building **REST APIs** with **FastAPI** and **Uvicorn**.
- Accelerate heavy image generation API throughput by **7×** while sustaining **100%** success under GPU bound workloads, leveraging **multiprocessing** (per model CUDA streams) and **multithreading** for **CPU/IO** tasks.
- Fine-tune **Whisper** for Arabic ASR, **Spark-TTS** for Arabic neural TTS, and **Qwen-3** for Arabic LLM agent workflows with **LangChain**, leveraging **Transformers**, **LoRA**, **PEFT**, and **Unsloth**.
- Deploy containerized APIs with **Docker** on cloud, implement auto scaling, custom logging to an OSS stack, and visualize metrics & alerts via **Power BI**.
- Lead a team of **3 junior engineers** with hands-on mentorship, code reviews, and shared best practice guidelines.

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