Abdulmalik Alquwayfili

📞 055 476 7376   ✉ af.alquwayfili@gmail.com   📍 Riyadh   in Abdulmalik Alquwayfili   ⦿ AbdulmalikDS

🔗 Portofolio

PROFILE

Driven to create real impact, I specialize in advanced deeptech solutions leveraging Natural Language Processing (NLP), Large Language Models (LLMs), and Agentic AI. I apply research-backed methods to design intelligent systems that simplify complex problems, enhance decision-making, and deliver measurable business value.

EDUCATION

Reinforcement Learning, CS234, *Stanford University*   Nov 2025

Bachelor of Information Systems, *Imam Mohammed Bin Saud Universty*   Sep 2025

EXPERIENCE

Artificial Intelligence Developer, *TAHAKOM • NLP Department* ⧉   Jul 2025 – Present

- Architected and deployed continuous **pre-training and post-training** pipelines on **AWS SageMaker Hyperpod**, leveraging Trainium, **GPU clusters**, and Hugging Face Optimum Neuron, covering **SFT, DPO, RLHF** for advanced alignment.

- Engineered a **parameter-efficient fine-tuning** pipeline for the **Qwen-2.5-72B** model using **LoRA** and **DeepSpeed ZeRO-3** on AWS A100, creating a **bilingual** instruction-following assistant specialized in **Saudi dialects**.

- Developed **LangGraph**-based **autonomous agents** with dynamic **tool calling** and real **code execution**, integrating **deep search** and file management. Collaborated with cross-functional teams including **Data Science, Product, and Video Understanding** on multiple use cases.

- Designed and **orchestrated 5+ autonomous agents** via **MCPs**, integrating ArXiv and GitHub APIs to **automate literature** review by **searching papers, code and data.**

- Conducted deep **literature reviews** on **agentic AI frameworks** (SICA, **LangGraph**, ReAct, Camel AI, Pydantic AI) and analyzed technical reports of state-of-the-art **LLM architecture**s and **alignment strategies**.

Machine learning Enginner, *Muragib* ⧉   Mar 2025 – May 2025

- **Trained, fine-tuned,** and evaluated 10+ open-source **BERT**-based models (e.g., CAMEL BERT, AraBERT), a**chieving over 90% accuracy** with <60ms latency, optimized for real-time applications.

- Developed and deployed advanced machine learning models for a wide range of **NLP tasks**, including **sentiment analysis**, sarcasm detection, location distribution analysis.

- Built scalable, **Al-powered analytics** systems that deliver actionable insights from unstructured text, empowering business stakeholders to make fast, data- driven decisions.

Research And Development Engineer, *وعى | Wae* ⧉   Nov 2024 – Mar 2025

- Designed, developed, and deployed Al-based solutions for hands-on workshops and internal tools, **directly impacting over 150 employees and 50+ students.**

- Contributed to the **design and delivery** of educational and technical content, bridging the gap between Al theory and practice in professional settings

- **Researched the latest Al tools** and technologies to drive innovation.

Software Developer - Artificial Intelligence VW, *Microsoft*   Jun 2024 – Jul 2024

- Built an AI-powered **computer vision model** using Custom Vision to **recognize and classify images**.

- Conducted **data preprocessing and feature engineerin**g for predictive analytics solutions.

SKILLS

Programming languages: Python, Java, JavaScript, SQL, Cuda, PHP, Rust, C++, R

Libraries, tools, and frameworks: Hugging Face, LLaMA-Factory, Unsloth, vLLM, LangGraph, LangChain, PyTorch, Keras, Scikit-learn, NLTK, OpenCV, Pandas, NumPy, Matplotlib, Seaborn, FastAPI, Ollama, Pydantic, wandb, Temporal

DevOps & Environments: Git, Bash, Linux, Ubuntu, Docker, Conda, Jupyter Notebook, Google Collaboration, Postman, Curls, Cursor, AWS SageMaker AI, HyperPod, HPC, SLURM, PyTourch Lightning, LogFire

Databases: PostgreSQL, MongoDB, Qdrant, ChromaDB, Metabase, Supabase, DBeaver, Alembic, Sqlalchemy

Cloud: Digital Ocean, AWS (SageMaker, Trainium), Alibaba Cloud, TPU, AI Foundry, Microsoft Azure

Model Architectures: Attention, Transformers, BERT, RoBERTa, DDP, SFT, DPO ,GRPO, RLHF, GPT, RAG, LoRA, Quantization, Agentic Al, Multi-Agent Systems, MCPS, MOE, MOA, RNNs EfficientNet, ConvNeXt

COURSES & CERTIFICATIONS

AWS AI & ML Scholars, *AWS & Udacity* ⧉   Jun 2025

Building Al Browser Agents, *AGI, inc* ⧉   Apr 2025

LLMOps: Building Real-World Applications With LLMs, *Udacity* ⧉   Dec 2024

Finetuning Large Language Models, *Lamini* ⧉   Dec 2024

Open Source Models, *Hugging Face* ⧉   Dec 2024

Neural Networks and Deep Learning, *DeepLearning.Al* ⧉   Sep 2024

Full Stack Deep Learning, *The Ful Stack*   Dec 2024

Python for Data Science, Al & Development, *IBM* ⧉   Jun 2024

PROJECTS

Quantigence: Multi-Agent Al Framework for Quantum Security   May 2025 – Jun 2025

Independent research on post-quantum cryptographic threats using a multi-agent system coordinating five quantized LLMS (Qwen2.5, Mistral, Llama3.1, Phi-3, Gemma2). Runs on consumer GPUs with Q4_K_M quantization and CUDA acceleration. Integrates data from Brave Search, arXiv, GitHub, Puppeteer, and SQLite using Multi-Context-Protocool (MCPs)

Farqad The Pioneer in Smart Financial Conversations, *GP* ⧉   Dec 2024 – May 2025

Farqad addresses the challenge of simplifying financial management and advice for Arabic and English-speaking users by developing a bilingual RAG chatbot with state of the art models, highly engineered and maintainable code

PerceiveAl, *Al League Finalist* ⧉   Feb 2025 – Apr 2025

This project aimed at making sports commentary accessible to the deaf and hard- of-hearing. By leveraging cutting-edge Al technologies, PerceiveAl translates live sports commentary into text thedn to sign language using animated avatars

langgraph-runtime-inmem-open, *LangGraph Contribution* ⧉   Jul 2025

open-source LangGraph runtime tools for agent memory management

Multi-Agent System Visualizer ⧉   Jul 2025 – Sep 2025

A sophisticated 3D visualization platform for multi-agent research systems with real-time AI-powered research capabilities.

Arabian Sheep Classification, *Kaggle Computer Vision Competition* ⧉   Jun 2025

Scored First in the leaderboard Day One and developed Al system for classifying 7 Arabian sheep breeds celebrating Eid al-Adha heritage, achieving championship-level 96.67% macro F1 score using ConvNeXt deep learning architecture with advanced data augmentation and Test-Time Augmentation techniques.

PUBLICATIONS

Quantigence: A Multi-Agent AI Framework for Quantum Security Research ⧉   Jun 2025