# HAMZA MOORAJ

Machine Learning Engineer — Multimodal & Agentic Systems | ML Systems
Pakistan | +92 317 3499268 | hhmooraj@gmail.com | Portfolio | GitHub | LinkedIn | HuggingFace |

*AI engineer focused on building and deploying multimodal and agentic systems for real-world applications. Experience designing end-to-end ML pipelines, integrating retrieval and model components, and optimising system performance under practical constraints.*

## EXPERIENCE

***AI "Learning Buddy" Project Intern — Heriot-Watt University***
*July 2025 – September 2025*

- Built an end-to-end retrieval-based learning assistant using LlamaIndex and Milvus.
- Implemented ingestion, chunking, and metadata-aware indexing pipelines for course materials.
- Deployed the system in a containerised environment and exposed a live demo for testing.
- Enables scalable academic support, reducing student search time and faculty query load.

***Data Science & Machine Learning Intern — Love For Data***
*August 2023*

- Built a data pipeline and popularity index across 1,970 Play Store apps (US and Pakistan datasets).
- Trained and evaluated a Random Forest model to identify key performance drivers.
- Enables cross-market analysis of app performance, informing feature prioritisation and market strategy.

***Software Engineering Intern — Folio3***
*June 2023 – July 2023*

- Designed and managed a relational database using MariaDB for a Visitor Management System.
- Integrated database with backend via PHP, enabling data flow across system components.
- Supports real-time visitor tracking, improving accuracy and reducing manual effort and overhead.

## PROJECTS

***A Systematic Exploration of Architectural Trade-offs for Crop Disease Classification - Lead Author***
*July 2025 – April 2026 — Accepted to Transactions on Machine Learning Research* ***(TLMR)****, 2026 — arXiv Preprint*

- Built a multi-source dataset pipeline (12 datasets → unified 65-class schema with lab/field splits).
- Implemented training workflows and evaluation pipelines for CNN and VLM architectures.
- Enables model selection for real-world deployment, supporting improved yields and reduced waste.

***Medical Question-Answering (QA) Conversational Agent (Course: F21CA) - Project Lead***
*January 2025 – April 2025*

- Built a RAG-based medical QA system over 411 NHS documents using FAISS retrieval.
- Designed a tri-modal pipeline integrating text, speech, and vision components.
- Enables grounded medical QA, improving reliability and trust in safety-critical applications.

***Logit-Weighted Suspicion Scoring in AI Control - Hackathon Project***
*February 2026 — Apart Research AI Control Hackathon — Submission*

- Built and evaluated an LLM-based monitoring pipeline for detecting suspicious agent behaviour.
- Implemented evaluation workflows using AUC-ROC to compare scoring strategies across monitors.
- Improved understanding of scoring limits in AI monitoring systems such as ControlArena.

## EDUCATION

***Heriot-Watt University — BSc Computer Science (Artificial Intelligence)***
*First Class Honours, 3x Deputy Principal's Award*
***BlueDot Impact — Technical AI Safety Course (2026) — Credentials***

## SKILLS

***Languages:*** Python, Java, C#, TypeScript, C
***Machine Learning:*** Transformers, Vision-Language Models, PEFT (LoRA), RAG
***Systems & Infrastructure:*** PyTorch, Milvus, Docker, Linux, CUDA
**Tools:** Git, Weights & Biases, RunPod, LangGraph