Muhammad Haroon

Islamabad, Pakistan | +92 348 8330334 | muhammadharoon8124@gmail.com
haroon-ai1.vercel.app | linkedin.com/in/haroon-ai | github.com/haroon-ai1 | huggingface.co/haroon8124

PERSONAL SUMMARY

AI Engineering student who builds and ships end-to-end ML systems: a RAG-based customer support chatbot (FAISS + LLM APIs, containerized with Docker), a visual search engine reaching mAP 0.93 (live on Hugging Face Spaces), and CNN classifiers. Strong foundation in PyTorch, retrieval systems, and rigorous model evaluation, with deployment experience across FastAPI, Streamlit, and Docker.

EDUCATION

SZABIST Islamabad Sep 2023 – June 2027

Bachelor of Science in Artificial Intelligence — CGPA: 3.5 / 4.0

Relevant Coursework: Deep Learning, Computer Vision, NLP, Machine Learning, Neural Networks, Data Mining, Database Systems, DSA, Operating Systems, Technical & Business Writing

WORK EXPERIENCE

CodeAlpha — Machine Learning Intern (Remote) July 2025 (1 month)

- Trained and evaluated supervised regression and tree-based models (Linear Regression, Decision Trees) on structured datasets; improved baseline R^2 and MSE on held-out validation splits through feature engineering and hyperparameter tuning.

PROJECTS

SupportGenie — RAG Customer Support Chatbot | Live Demo · GitHub

Python, FastAPI, LangChain, FAISS, Sentence-Transformers, Llama 3.3 (Groq API), Docker, Hugging Face Spaces

- Built a white-label RAG customer support platform that enables businesses to deploy AI assistants grounded in their own documentation, reducing hallucinations through retrieval-augmented generation and confidence-based human handoff.
- Developed an end-to-end knowledge pipeline including document ingestion, chunking, MiniLM embeddings, FAISS indexing, and LLM-based answer generation with source citations.
- Engineered a provider-agnostic LLM layer compatible with OpenAI APIs (Groq, DeepSeek, OpenRouter), calibrated retrieval confidence thresholds for escalation, achieved an 83% Hit@3 retrieval rate on a held-out evaluation set, and deployed the application using Docker on Hugging Face Spaces.

PixSearch — Content-Based Image Retrieval System | Live Demo · GitHub

Python, PyTorch, OpenCV, scikit-image, FAISS, Streamlit, Hugging Face Spaces

- Built a visual search engine retrieving the most similar images from a 1,000-image corpus using deep features (ResNet-50, CLIP ViT-B/32) against a classical baseline (HSV histogram + HOG + ORB), achieving mAP 0.93 on retrieval evaluation.
- Indexed L2-normalized embeddings in a FAISS flat inner-product index, achieving sub-millisecond top-K retrieval over the full corpus on CPU.

AgriVision-Pro — Plant Disease Detection System | GitHub

Python, TensorFlow/Keras, OpenCV, Image Processing

- Developed a computer-vision pipeline classifying crop leaf images into healthy vs. multiple disease categories (79% test accuracy) with resizing, normalization, and augmentation-based preprocessing.
- Reported per-class precision and recall alongside overall accuracy to surface minority-class weaknesses rather than relying on a single headline number.

Personal Finance Manager | GitHub

Python, TensorFlow/Keras (LSTM), Pandas, Matplotlib

- Developed a backend-driven financial management application supporting income/expense tracking, budgeting, loans, and recurring payments.
- Integrated LSTM models for expense forecasting (84% accuracy) and built visualizations for cash-flow trends and spending analysis.

CERTIFICATIONS

- **Machine Learning with Python — IBM (Coursera):** coursera.org/verify/PDLKU2ZZWKBL
- **Introduction to Data Science in Python — University of Michigan:** coursera.org/verify/WQ9KDVOTF0Z1
- **Machine Learning — DeepLearning.AI:** coursera.org/verify/JO7EIUHTUB4S

TECHNICAL SKILLS

Languages & Core: Python, C++, SQL, Git, DSA
ML / DL: PyTorch, TensorFlow/Keras, scikit-learn, NumPy, Pandas, CNNs, LSTMs, Transfer Learning, OpenCV, scikit-image
GenAI & Retrieval: RAG, LangChain, FAISS, LLM APIs (Groq / OpenAI-compatible)
Deployment: FastAPI, Docker, Streamlit, Hugging Face Spaces