Ahmed Abdullah

github.com/ahmedembeddedxx
business.ahmadabdullah@gmail.com
linkedin.com/ahmedembedded
+92-331-622-4618

Experience

- **Progyny Inc.** — New York, US (Remote)
  Data Analytics Engineer, Healthcare — *Jun 2025 – Present*
  - **ML and Data Platform Engineering**: Designed and scaled company-wide ETL and machine learning data pipelines using Apache Airflow, Marauder, and Dagster. Integrated Salesforce, Snowflake, and unstructured document sources using Python and SQL to support production ML and personalization workflows. Optimized schema mappings and transformation logic, reducing pipeline latency from 60+ seconds to approximately 3 seconds, achieving a 95% performance improvement and improving data reliability across internal tools and teams.
- **MacroMed (Pvt. Ltd.)** — Lahore, PK
  Machine Learning Engineer — *Feb 2025 – Jun 2025*
  - **Production Conversational AI and Recommendation Systems**: Led a team of three engineers to design and deploy a production-grade conversational AI platform using the RAFT framework. Built real-time recommendation pipelines backed by PostgreSQL, achieving sub-2 second end-to-end response latency and approximately 0.5 second recommendation retrieval. Deployed and scaled the full-stack system on DigitalOcean with a Next.js frontend, integrating ChatGPT-OSS and Mistral Nemo to deliver reliable, low-latency LLM inference in production.
- **Multimedia Mining and Search Group, Johannes Kepler University** — Linz, AT (Remote)
  Research Engineer, Multimodal Learning — *Aug 2024 – Feb 2025*
  - **Multimodal Representation Learning**: Developed deep learning models for face-voice association under missing modality conditions, improving accuracy by 12% on the VoxCeleb benchmark. Applied audio-visual architectures including WavLM, ECAPA, and VGGVox, and redesigned the training pipeline with reusable modules, reducing processing time by 10% and experiment setup time by 70%.
- **Machine Intelligence Group, FAST-NUCES** — Lahore, PK
  Research Engineer, Medical AI — *Aug 2024 – Nov 2024*
  - **Medical Imaging and Outcome Prediction**: Built an AI system to predict stroke outcomes from MRI scans, improving classification accuracy from 44% to 57% through enhanced image preprocessing, optimized augmentation pipelines, and integration of heterogeneous clinical data sources.

Education

- **National University of Computer and Emerging Sciences** — Lahore, Pakistan
  Bachelor of Science in Data Science; GPA 3.6 — *Aug. 2022 – May. 2026*
  - **Honors**: Dean's List of Honor, Top of Batch
  - **Experience**: Teaching Assistant for OOP (CS1004) and COAL (EE2003)

Projects

- **CrickAI – LBW Review System using YOLO**: Built an automated video review system for LBW decisions in cricket using YOLO for ball and player detection. Curated and annotated datasets using Roboflow and LabelImg, and optimized detection accuracy for real-time decision support in sports analytics workflows.
- **PolyHope-M – Hope Speech Classification**: Developed a multilingual text classification model using LLaMA and PyTorch to identify supportive speech in online content, achieving 1st place among 400+ submissions on CodaLab with top-ranked benchmark accuracy.
- **DermaScan3D – Skin Cancer Detection using CNNs**: Built a multimodal CNN-based skin cancer classification system using TensorFlow and ResNet architectures, ranking in the top 1% among 8000+ participants on Kaggle through optimized preprocessing and model tuning.

Programming Skills

- **Languages**: Python, SQL, C++.
- **Technologies**: PyTorch, TensorFlow, Hugging Face Transformers, OpenCV, Scikit-learn, MONAI.
- **Tools & Platforms**: LangChain, Vector Databases (FAISS), Apache Airflow, Snowflake, PostgreSQL, Docker, MLflow.