

Afsah ur Rehman

 GitHub  afsahurrehman11@gmail.com  +923264565399  LinkedIn

EDUCATION

Information Technology University (ITU)

BS Computer Science

Lahore, Pakistan

SKILLS

Python | PyTorch | MERN Stack | Deep Learning | Data Engineering | Cloud (Azure, AWS) |
Machine learning | Azure DevOps | React | Angular | MYSQL | Node Js | MongoDB | Git & Github

PROFESSIONAL EXPERIENCE

Data Engineering Intern – AI Focus

Jul 2024 – Aug 2024

DigiFloat

- Built ETL pipelines with Azure Data Factory, Databricks, and Functions; transformed data in Blob Storage, Data Lake Gen2, and Azure SQL; queried via Synapse; fine-tuned and deployed GPT models; set up CI/CD with Azure DevOps and deployed web apps.
- Tools: *Azure DevOps, Data Factory, Databricks, Functions, Synapse, Blob Storage, Data Lake, Azure SQL, GPT API, Python, PySpark, SQL, Pandas, NumPy*

Web Development Intern

Jun 2023 – Aug 2023

TIERS Limited

- Build blog, Ecommerce, Portfolio, University websites
- Tools: *React, MySql, Node Js, Bootstrap, JavaScript*

PROJECTS

FlowScript: Video Narrative Generator | AI/ML

Built AI system for video marketing narratives, capturing audio, visuals, text, 3D motion. Used Whisper for audio with noise & music filtering, CLIP for frame selection, BLIP for captions, Flan-T5 for motion analysis, EasyOCR for text. Merged data for Flan-T5 to create 3 narratives, CLIP for visual precision. Added hashtags, contact. (*Python, Web app*)

PulseForge AI: Medical Tabular Data Synthesis | AI/ML

Generated realistic synthetic medical records with user-configurable anomaly rates; built a Python API and trained 4 generative models (CTGAN, TVAE, Tabular Diffusion, Masked Autoregressive Flow) plus 2 outlier samplers (Bayesian Network, Random Forest). (*Python, PyTorch, scikit-learn*)

DocLens-AI

Built a Python-based RAG agent that ingests PDFs, embeds text with HuggingFace, indexes with FAISS, and queries a Groq LLM to deliver precise, context-aware answers. (*Python, LangChain, FAISS, Groq-Client*)

AirKeys | AI/ML

Implemented touchless HCI (hands-free mouse & keyboard system): left index for cursor & pinch-click; right index for hover & key input; right-index swipe any direction for arrow navigation. (*Python, OpenCV, MediaPipe, PyAutoGUI*)

Uno Cards Classifier | AI/ML

Implemented a custom ANN from scratch with advanced augmentation and tuning; reached 99.5% on a self-collected dataset and ~70% on external images. (*Python, NumPy, OpenCV*)

CIFAR-10 Classification Trifecta | AI/ML

Built PCA-based, autoencoder, and CNN classifiers; reduced PCA training time by 98%, extracted 64-dim embeddings, and achieved top CNN accuracy with minimal overfitting. (*Python, NumPy, PyTorch*)

NLP TF-IDF Search Pipeline | AI/ML

Designed an end-to-end TF-IDF engine over 21,941 ACL papers for instant top-5 query retrieval. (*Python, NLTK, NumPy*)

High-Dimensional Regression & Neural Nets | AI/ML

Applied linear regression via normal equations, gradient descent; built a neural net for logical gates like AND, OR. (*Python, Pandas, NumPy, scikit-learn*)

Multi-Tier VPC Deployment | Cloud

Provisioned an AWS VPC with a public app subnet (via Internet Gateway) and a private DB subnet (via NAT Gateway); enforced routing, Security Groups, and NACLs for strict traffic control. (*AWS VPC, IGW, NAT Gateway, Route Tables, Security Groups, NACLs, EC2*)

ETL Pipeline Creation | Cloud

Built ELT workflows using Data Factory, Azure Functions, Blob & Data Lake Gen2, and Azure SQL Database. (*Azure Data Factory, Azure Functions, Blob Storage, Data Lake Gen2, Azure SQL*)

CERTIFICATES

AWS re/Start Program Graduate

Completed 12-week **AWS Cloud Practitioner** curriculum; awarded AWS Cloud Practitioner exam voucher and AWS re/Start graduate badge (exam scheduled).