Aimal Rehman

Atlanta, GA (open to relocation)
(470) 338-7353 rehman.aimal@gmail.com aimalrehman aimalrehman92.github.io aimalrehman92

Education

Georgia State University **Sep 2025 – Present**
MS Applied Statistics - Machine Learning & Data Science (4.00/4.00) *Atlanta, GA, USA*

Lahore University of Management Sciences (LUMS) **Sep 2017 – Jun 2019**
MS Electrical Engineering (3.22/4.00) *Lahore, Pakistan*

COMSATS University Islamabad **Feb 2012 – Jan 2016**
BS Electrical Engineering (3.27/4.00) *Lahore, Pakistan*

Experience

Turing.com **May 2024 – Aug 2025**
Data Scientist (Remote) *Palo Alto, CA, USA*

- Led the development of high-quality Python and SQL training data for Google Gemini, covering end-to-end data analysis, machine learning, and production pipeline workflows
- Created post-training data across SFT for teaching desired solutions, RM for learning human preferences, and RLHF for optimizing model behavior
- Fine-tuned and evaluated internal Gemini models for coding accuracy, analytical reasoning, instruction-following, and solution quality

Tkxel **Feb 2024 – Jul 2024**
Senior Data Scientist & Tech Lead *Lahore, Pakistan*

- Led the development of a Llama-3-based taxation chatbot for a European tax-filing company, helping individuals and businesses navigate the platform and complete tax returns
- Designed RAG and data-ingestion pipelines for text and tables across PDF, DOCX, CSV, and Excel files; successful delivery generated an additional $0.1 million in revenue
- Led an internal Executive Assistant GPT with company Q&A, sales-deck generation, and case-study creation using HR, marketing, and sales data
- Implemented vector search, multi-step function calling, and session-level conversational memory to support reliable, context-aware interactions

Clinical & Translational Imaging Lab (LUMS) **Jul 2022 – Jun 2023**
Research Associate (Contract) *Lahore, Pakistan*

- Developed and evaluated U-Net, Conditional GAN for cardiac MRI synthesis and semantic segmentation
- Built deep learning pipelines in PyTorch and TensorFlow to segment the right ventricular blood pool, left ventricular blood pool, and myocardium
- Developed a novel 3D–2D Hybrid U-Net for robust segmentation from lossy MRI scans, enabling cardiac-volume estimation across systolic and diastolic phases
- Supported more consistent assessment of cardiac function and disease-related abnormalities for radiologists and clinicians

TG Technologies (TechGenies) **Jan 2022 – Jan 2024**
Data Scientist (Remote) *Islamabad, Pakistan*

- Analyzed workforce data across multiple organizations to identify patterns in employee retention, absenteeism, performance, and staffing
- Developed absenteeism forecasting pipelines using ARIMA for short-term prediction and Facebook Prophet for longer-term workforce planning
- Built interpretable employee churn models using Random Forest and Gradient Boosting to identify key drivers of attrition
- Applied OCR and NLP to develop a resume–job matching product that automated candidate shortlisting and reduced manual screening effort
- Converted HR analytics proofs of concept into funded client engagements and revenue-generating products and maintenance services

Afiniti **Aug 2020 – Nov 2021**

Data Scientist I *Islamabad, Pakistan*

- Analyzed client contact-center data and produced integrity reports to identify data-quality issues affecting model performance
- Trained and deployed Decision Tree and STAN-based models for real-time caller–agent pairing in production contact centers
- Monitored deployed models, evaluated performance, and performed feature engineering and selection using Random Forest and LightGBM
- Optimized predictive pipelines through grid search, contributing approximately £90K in monthly revenue gain on Sky UK's largest queue

Information Technology University **Sep 2019 – Jan 2020**

Teaching Assistant (Contract) *Lahore, Pakistan*

- Taught the graduate course Statistical & Mathematical Methods for Data Science to the MS Data Science class as well as prepared & evaluated grading instruments

Electronics & Embedded Systems Lab (LUMS) **Mar 2016 – Mar 2017**

Research Assistant (Contract) *Lahore, Pakistan*

- Designed a hardware-efficient NOMA receiver for iterative multiuser detection and channel decoding in non-orthogonal communication systems
- Applied Maximum Likelihood, MAP estimation, and Monte Carlo sampling to optimize belief-propagation decoding over bipartite factor graphs
- Improved transmitted-signal recovery while minimizing computational complexity and transmission energy for resource-constrained receiver architectures

Technical Skills

Languages: Python, R, SQL, C

ML & Data Frameworks: PyTorch, TensorFlow, Scikit-learn, PySpark, OpenCV, Scikit-LLM

Generative AI & Application Frameworks: LangGraph, Google ADK, LlamaIndex, CrewAI, Pydantic AI, OpenAI, Ollama, Streamlit, Gradio, Flask, FastAPI

Data Tools & Platforms: JupyterLab, VS Code, DBeaver, Tableau, Power BI

Cloud & DevOps: AWS (S3, SageMaker, Bedrock, Lambda), Docker, Git, Linux

Research

- ragR: Retrieval-Augmented Generation and RAG Assessment in R - 2026 [arXiv preprint]
- Towards design and automation of hardware-friendly NOMA receiver with iterative multi-user detection - 2017 [Link]

Certificates

- Agentic AI from deeplearning.ai
- 5-Day AI Agents Intensive Course with Google/Kaggle
- Basic HIPAA Privacy & Security Certification

Others

- Served as a volunteer with Boy With a Ball in Atlanta, GA, 2026
- Served as Boys Hostel Proctor, COMSATS University Islamabad, 2015
- Served as an active member of the Civic Engagement Society, COMSATS University Islamabad, 2015