

# MUHAMMAD SHAHEER JAMAL

Karachi, Pakistan • Shaheer-jamal@outlook.com • [LinkedIn](#) • [Github](#) • 03323501258 • [Website](#)

---

## WORK EXPERIENCE

### Data Science Intern @ 10Pearls

Dec 25 - Present

- Architecting serverless end-to-end ML pipeline for real-time AQI forecasting, integrating external APIs for automated data ingestion and implementing preprocessing workflows for time-series environmental data
- Developing multi-model training framework with Hopsworks feature store for experiment tracking and version control
- Building continuous retraining pipeline using GitHub Actions triggered by new API data, enabling automated model updates and deployment in serverless production environment

### ML Intern @ Resolve

Jun 25 - Aug 25

- Built legal document querying system using RAG (Retrieval Augmented Generation) architecture with hybrid retrieval combining semantic and keyword search
- Engineered document processing pipeline handling 300+ pages legal documents using Langchain and ChromaDB vector database
- Achieved 85%+ accuracy in legal query responses through prompt engineering and retrieval optimization
- **Skills** : Langchain, Ollama, RAG, ChromaDB, NLP

### AI Intern @ Arcana Info Pvt Ltd

Aug 24 - Oct 24

- Developed chronic kidney disease prediction models using IBM watsonx.ai and scikit-learn, implementing end-to-end ML pipeline from data preprocessing to model evaluation
  - Created interactive business intelligence dashboards using IBM Cognos Analytics to visualize healthcare data trends and model performance metrics
- 

## PROJECTS

### Heart Disease Prediction - MLOps Pipeline

- Built end-to-end ML pipeline with XGBoost for cardiovascular risk prediction, implementing automated retraining workflows using GitHub Actions for continuous model updates
- Engineered feature store integration with Hopsworks for versioned data management and batch inference pipeline for efficient model deployment
- Deployed production-ready Streamlit application with real-time predictions, incorporating user data collection for model improvement cycles

### NutriScan - Food Recognition & Nutrition Analysis System

- Developed full-stack computer vision application using PyTorch for food image classification across 5+ Pakistani food categories
  - Deployed custom CNN model on Azure Machine Learning with REST API integration, achieving 80%+ classification accuracy on test dataset
  - Built scalable backend with MongoDB for nutrition data storage and Streamlit frontend for real-time image processing
- 

## EDUCATION

### Bachelor of Science in Artificial Intelligence

2022 - 2026

Dawood University of Engineering and Technology

- **GPA**: 3.7/4
- 

## SKILLS

- **Languages & Frameworks**: Python, C, R, SQL, Assembly Language, PyTorch, Scikit-learn, TensorFlow
  - **AI/ML Specialization**: Computer Vision, NLP, Deep Learning, RAG Architecture, Model Deployment
  - **LLM Technologies**: Langchain, Ollama, Document Embedding, Vector Databases (ChromaDB), AI Agents
  - **Cloud & DevOps**: Microsoft Azure (Azure ML, App Services), REST APIs, Git, Docker basics
  - **Databases**: MongoDB, PostgreSQL, ChromaDB, Vector Databases
- 

## CERTIFICATIONS

- **Microsoft Certified: Azure AI Fundamentals (AI-900)**
- Machine Learning Specialization from **DeepLearning.ai** and **Stanford Online**
- Neural Networks and Deep Learning from **DeepLearning.ai** and **Stanford Online**