

King Abdullah Bin Abdulaziz Rd., 12445, Riyadh

□ (+966) 53-9935-783 | Sahmad786.writes@gmail.com | Aahmad786writes.github.io | Inahmad786writes | Inahmad786-writes

"Between noise and pattern, I find meaning. Between data and code, I build purpose."

# Summary.

Machine Learning & Generative AI Engineer based in Riyadh with 3+ years of experience delivering production-grade solutions using LLMs, Diffusion Models, and VAEs. Skilled in fine-tuning models like GPT, LLaMA, and Mistral for finance, automation, and customer service. Proficient in RAG pipelines, vector databases, and agentic frameworks (LangChain, CrewAl, Hugging Face). Experienced in QLoRA, quantization, and deploying scalable solutions on AWS and GCP to drive real-world business impact.

# **Education**

### **GIFT University**

Gujranwala, Pakistan

BACHELOR'S OF SCIENCE IN COMPUTER SCIENCE

2019 - 2023

- Related Coursework: Machine Learning, Artificial Intelligence, Natural Language Processing, Data Structures and Algorithms, Database Systems.
- First Position in Final Year Project, received funding from Ignite for project development

# Work Experience \_\_\_\_\_

**GAIT** Riyadh, Saudi Arabia

AGENTIC AI ENGINEER - REMOTE

Mar. 2025 - Apr. 2025

- Designed and deployed autonomous AI agents for KnowYour.AI, automating company intelligence workflows, reduced manual research time by 85% and handled data for 1,000+ enterprises.
- Developed an LLM-integrated chatbot capable of retrieving structured data from Google Sheets using natural language prompts, streamlining access to 10K+ records with a 70% faster response time.
- Built AI agents for Watch2Give, a gamified impact platform, including systems for proof validation and reward dispatch using LangGraph and vision models — cut human review time by 60% and improved accuracy in early testing.
- · Automated proof verification and staking optimization using LLM + vision agents, accelerating validation time by 60% and increasing vendor participation by 40% during MVP rollout.

**Devster Labs** Islamabad, Pakistan

ARTIFICIAL INTELLIGENCE ENGINEER

May. 2024 - Nov. 2024

- Developed and deployed AI chatbots that automated 70% of routine support queries, leading to a 15% boost in response efficiency and reducing customer support costs by 10%.
- Fine-tuned GPT-4 using QLoRA for domain-specific tasks; integrated with vector DBs (FAISS) for fast retrieval in RAG pipelines.
- Implemented CI/CD pipelines for ML models, enabling automated retraining and real-time performance monitoring with MLflow/Kubeflow.
- · Optimized transformer models with quantization and multi-GPU on GCP, reducing latency by 15%.
- Optimized audio processing models for Sound Event Detection Systems, reducing processing time by 3%.
- Deployed Al-driven weapon and anomaly detection systems in real-world environments, leading to a 40% reduction in violencerelated incidents.

**GIFT University** Gujranwala, Pakistan

#### RESEARCH ASSISTANT & TENSORFLOW DEVELOPER

Feb. 2023 - Mar. 2025

- Trained GAN sketch-to-image model on GPU, optimized for CPU/GPU deployment; achieved 80% accuracy beating benchmarks.
- Built risk prediction models with 85% accuracy, enabling early detection of defaults in academic finance studies.
- Developed a deep learning model using CNNs to analyze satellite imagery, achieving 85% accuracy in identifying poverty-stricken regions, and aiding in resource allocation for underserved areas, which helps targeted support and policy planning.
- · Engineered and optimized zero-shot, few-shot, and chain-of-thought prompts for LLMs, as demonstrated in projects like TailorMade-Al and Multi-agent-Customer-Support-Automation, enhancing Al-driven decision-making and user interactions.

CodSoft Kolkata, India

#### MACHINE LEARNING INTERN - REMOTE

Oct. 2023 - Nov. 2023

Used ResNet-based encoder and custom decoder for image captioning and explored YOLOv5 for object grounding.

- Boosted recommendation CTR by 25% via collaborative filtering, improving engagement on deployed platforms.
- Integrated machine learning models like ChatGPT, Dall-E, Gemini and others into real-world applications using Flask REST APIs for seamless web application interaction.
- Optimized data pipelines and preprocessing for sound event detection, improving real-time classification accuracy on GCP.

**Publications** 

# 25th International Conference on Digital Image Computing: Techniques and Applications

Perth, Australia

LOCALLY-FOCUSED FACE REPRESENTATION FOR SKETCH-TO-IMAGE GENERATION USING

Noise-Induced Refinement

2024

• This paper introduces a novel approach for sketch-to-image generation, focusing on localized face representation with noise-induced refinement to enhance visual accuracy and realism.

# International Conference on Innovation in Artificial Intelligence and Internet of Things

Jeddah, Saudi Arabia

LOCALLY FOCUSED MULTI-LEVEL FEATURES-BASED FRAMEWORK FOR POVERTY ESTIMATION

Submitted for Publication 2025

• This study proposes a novel multi-level feature extraction, fusion, and dual-attention mechanism framework for poverty estimation via satellite imagery.

# 3rd International Conference on Energy, Power, Environment, Control and Computing

Gujranwala, Pakistan

SKETCH TO IMAGE SYNTHESE: HARNESSING DEEP LEARNING TECHNIQUES FOR REALISTIC VISUAL

2025

TRANSFORMATION

Proposed a deep learning solution to transform black-and-white face sketches into realistic color images, addressing challenges
in suspect identification. Submitted this work for AI applications in law enforcement.

# **Projects**\_

#### **Human Sketch-to-Face Conversion**

2023

- Tech used: Python, Pytorch, GANs, CNN, Autoencoders
- Built a GAN-Autoencoder to convert sketches into realistic faces.

### **RAG-based Conversational AI Chatbot for Financial Advisory**

2024

- Tech: Python, TensorFlow, Hugging Face, LangChain, OpenAl, CrewAl, AutoGen
- Created a highly accurate chatbot using transformer-based models like GPT-3 and BERT.
- Developed a RAG-based chatbot for fintech, enhancing support efficiency, transferable to insurance customer service.

#### **KnowYour.AI Company Intelligence Platform**

2025

- Tech used: Python, CrewAl, LangChain, Groq, Chain-Of-Thought Prompting, FAISS
- Built CrewAl agents for autonomous company profiling and summarization, powering KnowYour.Al with a cross-functional team.

### Watch2Give - Al-Powered DeFi Ecosystem

- Tech used: LangGraph, OpenAI, Groq, Few Shot Prompting
- Developed LangGraph agents for donation routing, staking logic, and image validation in a gamified cross-chain platform.

#### **Sound Event Detection System**

2024

- Tech used: Python, TensorFlow, Keras, Librosa
- Built a deep learning-based sound event detection system with 7% improvement in accuracy and reduced latency by 2% in realtime detection. Reduced model inference latency by 30% using quantization and model distillation.

# Agentic AI Systems - Real-World Problem Solving with Multi-Agent Architectures

2025

- Tech used: CrewAl, LangChain, LangGraph, OpenAl, Groq, FAISS, Vision Models
- Designed and deployed autonomous agents for tasks like company profiling, document parsing, and spreadsheet-driven data querying, handling 1K+ entities and reducing manual work by up to 85%.
- Built Al agent workflows with LangGraph and LLMs for vendor verification and proof validation in a gamified giving platform, improving task speed by 60% and boosting transparency.
- Focused on real-world impact: automating research, streamlining data access, and supporting underserved communities through intelligent agent flows.

Other Projects 2024

- Tech used: Python, Sci-kit learn, Pandas, Numpy, PyTorch, TensorFlow, CNN, RNN, SQL, Plotly
- Developed and implemented a variety of machine learning models, including:
  - Image Segmentation using U-net, FCN-8 architecture
  - Built a lightweight keyword spotting model using TensorFlow Lite for Microcontrollers.
  - Recommendation systems for personalized content delivery, resulting in a 5% enhancement in recommendation accuracy.
  - Time-series forecasting for stock price prediction.
  - Transformer Models for Language Translation.

May 22, 2025 Ahmad Liaqat · Résumé 2

# **Skills & Interests**

**Programming Languages** Python, Java, C, C#, SQL, LaTeX

ML and DL Algorithms Supervised, Unsupervised, Reinforcement, CNN, RNN, LSTM, Transformers, GAN, Autoencoders

Natural Language Processing Tokenization, POS tagging, NER, Classification, Word2Vec, BERT/GPT, Hugging Face, SpaCy, NLTK Computer Vision Image Classification, Object Detection, Face Recognition, Sketch-to-Image, OpenCV, MediaPipe

**Data Engineering** Data Cleaning, Preprocessing, Feature Engineering, SQL, Pandas, NumPy, MLflow, Kubeflow, ETL

**Frameworks & Libraries** PyTorch, TensorFlow, Keras, Scikit-learn, FastAI, CrewAI, LangChain, llamaindex

**Development Tools** Git, Docker, Jupyter Notebook, Google Colab, VS Code, REST APIs, Flask

**Evaluation & Visualization** F1-Score, Accuracy, Precision, Recall, Confusion Matrix, Matplotlib, Seaborn, Plotly, Power BI

**Cloud & MLOps** AWS, GCP, Vertex AI, BigQueryML, AutoML, MLflow, CI/CD

**Soft Skills** Problem Solving, Communication, Teamwork, Adaptability, Quick Learner, Attention to Detail

**Languages** English (Fluent), Urdu (Fluent), Arabic (Basic – currently improving for professional use)

## **Licenses & Certifications**

Ongoing IBM Data Science Professional Certificate, by IBM Skills Network Team		Coursera
2025	MIT 6.S191: Deep Learning Specialization, by Alexander Amini & Ava Amini	MIT
2025	TensorFlow: Advanced Techniques Specialization, by Laurence Moroney	DeepLearning.Al
2024	TensorFlow Developer Professional Certificate, by Laurence Moroney	DeepLearning.Al
2024	Natural Language Processing Specialization, by Younes Bensouda Mourri	DeepLearning.Al
2023	<b>2nd Runner Up: Al Driven Hackathon</b> , PAK-UK Academic Bridge and hosted by Comsats University	Pakistan
2022	Machine Learning Specialization on Google Cloud, by Prof. Andrew NG.	Coursera