



KHIZAR SULTAN, M.Sc.

(Senior AI/ML Engineer - Generative AI)

[in LinkedIn](#) | [+92-342-0076652](#) | [Portfolio](#) | [khizarsultan.datascientist@gmail.com](#) | [Lahore, Pakistan](#)

Professional Summary

Accomplished AI/ML Engineer with **6+ years** of experience and a **Master's** in Data Science. Delivered **7+ industrial** AI/ML projects across telecom, healthcare, insurance, retail, and cybersecurity. Proficient in **Python**, **Generative AI**, **Agentic AI**, **Chatbots** and **LLMs**. **2X Microsoft-certified** in Azure AI and Data Fundamentals with **20+ open-source** projects on GitHub showcasing AI/ML expertise.

Work Experience

Sr. AI/ML Engineer - Gen AI

NETSOL Technologies

Pakistan 09/2024 - Current

- Built a conversational **chatbot** for a car lease company using Generative AI, LLMs, **RAG**, and **LangChain**, and deployed it via **FastAPI**, **Docker**, and **AWS EC2**, which increased car sales by **15%** post-deployment
- Developed a RAG-based chatbot for teachers using LangChain, **LangGraph**, and **Agentic AI**, enabling teachers to query assignment submissions and automatically email students who had not submitted, **saving 5 hours** per day for each teacher.
- Designed a **Model Context Protocol** (MCP) server using **GofastMCP**, integrated with **OpenAI** and **Claude Desktop**, implementing intelligent tool calls and counter-questioning to enhance contextual accuracy in user interactions.
- Architected an **API gateway** on **Databricks** leveraging **Mosaic AI**, which routed user queries to the most relevant LLM automatically, streamlining API usage and reducing manual overhead.

AI/ML Engineer - Gen AI

SlashNext Inc

Pakistan 07/2022 - 09/2024

- Fine-tuned the **Large Language Models** including **Llama 2** for domain specific data generation, tuned the hyper-parameters, apply chain of thought **prompt engineering** and achieved **99%** accuracy in malicious sms generation.
- Led the development and implementation of deep learning **BERT** model for sms text classification, achieved an accuracy of **98%**, resulting in a **40% decrease** in **cyber security attacks**.
- Designed and implemented complete **pipeline** for **text processing**, model training, evaluation, and **deployment** using CI/CD pipeline, leading to a **25% reduction** in training, testing and deployment time.
- Monitor the performance of **deployed models**, implement new techniques to improve, **communicate findings**, progress, and results to **stakeholders** through email, confluence, reports and presentations.
- Work closely with lead data scientists and stakeholders to **understand requirements** and translate them into technical solutions, develop and **integrate** new **AI/NLP models** into existing production systems.

Data Scientist - AI/ML Engineer

Retailo Technologies

Pakistan 12/2021 - 06/2022

- Developed an **AI recommendation engine** utilizing models like Collaborative Filtering algorithm, achieving a revenue of \$1 million.
- Extracted sales data from **AWS Redshift** using **Python**, applying cleaning and feature engineering techniques with **Pandas** and **NumPy**, enhancing data quality and enabling more accurate predictions with models like **Random Forests** and Gradient Boosting Machines (GBMs).
- Optimized data storage systems, implementing techniques such as data compression and indexing, reducing access times by **40%**.
- Calculated KPIs and established a tracking pipeline to assess the impact of the recommendation engine on **KPIs**.

Associate Data Scientist - AI/ML Engineer

Addo.ai

Pakistan 10/2020 - 12/2021

- Collaborated with international clients such as Singtel, and Comfortdelgro, delivering tailored solutions that boosted **efficiency by 15%**.
- Actively participating in daily scrum meetings to ensure timely project completion and reduce project turnaround **time by 10%**.
- Utilized Python libraries like easyocr, numpy, pandas, matplotlib, and **scikit-learn**, contributing to a **25% reduction** in data processing time.
- Trained and validated ML models, leading to a **30% improvement** in model accuracy, and deployed them using advanced techniques like **Docker** and Kubernetes which decreased deployment time by **20%**.

Education

M.Sc. Data Science

University of Management and Technology

Pakistan 07/2021 - 07/2023

- Major Courses: Artificial Intelligence | Natural Language Processing | Data Science | Machine Learning | Deep Learning | Computer Vision

B.Sc. Computer Science

National University of Computer and Emerging Sciences

Pakistan 07/2016 - 07/2020

- Major Courses: Data Structures and Algorithm | Programming | Statistics and Probability | Calculus | Linear Algebra | Data Mining

Skills

- **Generative AI:** Langchain, Model Context Protocol, Agentic AI, LLMs, RAG, Chatbots, n8n, Fast API, BERT, Llama, CAG, Prompt Engineering, Eval, Mosaic AI, Conversational Chatbot, FAISS, Milvus, Vector Databases, OpenAI, Claude, Gemini, Grok
- **Machine Learning:** Model Development, Model Deployment, Python, Transformers, TensorFlow, PyTorch, Pandas, Keras, NumPy, Docker, Flask, Fast API, AWS, SageMaker, S3, Azure Machine Learning, Databricks Apps, MLOps, Data Drift, Concept Drift
- **Leadership Qualities:** Problem-solving, Collaboration, Verbal and Written Communication, Teamwork, Conflict Resolution, Innovation.
- **Languages:** English (Professional)

Industry Projects

- **Enterprise Conversational RAG Chatbot:**

- **Objective:** Built a conversational AI chatbot to handle complex business queries and automate workflows for end-users.
- **Approach:** Utilized LLMs, RAG, LangChain, LangGraph, and Agentic AI to design context-aware interactions, integrated task automation (e.g., sending emails, retrieving student data), and deployed via **FastAPI + Docker** on AWS EC2 for scalability.
- **Impact:** Increased user engagement by **30%** and streamlined routine operations, enabling businesses to reduce manual effort while improving overall efficiency.
- **Malicious and Benign SMS Text Classification using Natural Language Processing:**
 - **Objective:** Develop a machine learning model capable of distinguishing between **malicious and benign SMS** texts to enhance mobile security and protect users from **phishing attacks** and malware threats.
 - **Approach:** Utilize **Natural Language Processing** techniques such as **word embeddings** to represent SMS texts as numerical features. Implement **BERT** deep learning model for text classification. Train the model on a labeled dataset containing examples of both malicious and benign SMS.
 - **Impact:** Leverage NLP techniques for spam detection. Achieve **98% training** and **97% testing** accuracy with BERT. Enhance mobile user security by **40%**.
- **Malicious and Benign Email Classification using Machine Learning:**
 - **Objective:** Develop a machine learning-based email classification system to detect and classify emails as **malicious** or benign, improving **cybersecurity** by identifying potential **threats**.
 - **Approach:** Employ Natural Language Processing techniques such as TF-IDF, **Word2Vec**, to extract features from email content (e.g., subject line, body text, sender information) and classify emails as malicious or benign. Implement anomaly detection models, such as **Random Forest** to identify unusual email patterns and flag potential malicious activities, like **phishing or spam**.
 - **Impact:** This implementation reduces the number of malicious emails reaching users by **40%**, improving overall **email security** and preventing **phishing attacks** and spam. Additionally, it enhances the accuracy of email classification, achieving a **95%** classification accuracy rate, which minimizes the time and effort required for manual email filtering.
- **Model Context Protocol (MCP) Server for Intelligent Tool Orchestration:**
 - **Objective:** Designed an MCP server to enhance multi-model AI systems with intelligent tool calling and contextual reasoning.
 - **Approach:** Implemented GoFastMCP, integrated with OpenAI and Claude Desktop, enabling dynamic tool calls based on user intent and counter-questioning for improved accuracy. Designed a modular architecture to support future integrations.
 - **Impact:** Improved response accuracy and reduced manual tool invocation, enhancing system adaptability and creating a scalable foundation for enterprise-grade GenAI solutions.
- **Powered Automated Identity Card Information Extraction System:**
 - **Objective:** Develop a system capable of **accurately extracting** name, date of birth, address, and identification numbers from national identity cards using **Optical Character Recognition (OCR)** technology.
 - **Approach:** The approach involves preprocessing the images of identity cards to enhance readability, and apply **deep learning** based OCR algorithms to **recognize and extract the** text using easyocr python library.
 - **Impact:** This project simplifies the process of digitizing information from identity cards, enabling **40% faster** and more efficient identity verification processes in various domains such as government services, banking, and travel.

Awards

- Recipient of the Dean's List Award (NUCES)
- Arctic Code Vault Contributor (Github)

Certifications and MOOCs

- Microsoft Azure **Certified Artificial Intelligence** Fundamentals
- Microsoft Azure **Certified Data** Fundamentals
- Machine Learning Specialization - University of Washington
- MLOps Specialization - Deeplearning.ai
- Deep learning Specialization - Deeplearning.ai
- Data Science Specialization - University of Michigan