

MOHAMMED AHMED PARACHA

+1 (647) 242-9068 | +92 333 3145702 | map.ai@outlook.com | linkedin.com/in/ParachaAI | www.ahmedp.com

SUMMARY

ML, Gen AI, and Forward Deployed Engineer with 10 years of experience deploying, and operating production-grade AI/ML systems in cloud environments and in-house servers, often embedded directly with enterprise clients to take solutions from prototype to production. Most recently led the AI team at Utopia Brands, building enterprise-wide GenAI automation and scalable inference infrastructure. Hands-on experience delivering end-to-end RAG pipelines and agentic AI workflows with LangChain and LangGraph. Strong Python engineering foundation with deep knowledge of model serving, observability, CI/CD automation, and AI platform governance.

EXPERIENCE

Director of AI | Utopia Brands

Sep 2025 - May 2026

- Established and led a team of AI engineers responsible for building and shipping production AI systems across the organization, from model development and serving through deployment, monitoring, and iteration.
- Developed a POC a proprietary Amazon Campaign Management System that automates active campaigns, with backend infrastructure handling high-frequency data ingestion and real-time bid adjustments.
- Led development of automation tools for Product Management teams and Media teams, streamlining cross-functional workflows through AI-powered internal tooling and reducing manual effort across content, listing, and campaign operations.
- Led Generative Engine Optimization (GEO) and Answer Engine Optimization (AEO) initiatives for Utopia Brands websites and Amazon listings to maximize AI-driven search visibility and product discovery.
- Reverse engineered a Fabric AI Inspection Machine to perform automated quality assurance on manufactured cloth, integrating computer vision into production workflows.

Principal Data Science Consultant | Systems Limited

Jan 2025 - Sep 2025

- Led end-to-end engineering of an enterprise-grade Retrieval-Augmented Generation (RAG) system with a ReAct agent chatbot for the Ministry of Tourism Saudi Arabia. Integrated LangChain, and LangGraph, for orchestration, tracing, and evaluation, with Neo4j Knowledge Graphs and Azure AI Foundry powering the retrieval pipeline.
- Engineered and deployed AI-powered FastAPI microservices, including Vision-Language Models and Face Recognition, into production environments serving 1,000+ internal users with high availability and real-time inference.
- Established robust MLOps and LLMOps pipelines using Azure DevOps and Docker, covering CI/CD automation, model versioning, testing, observability, and lifecycle management for GenAI systems in production.

Senior Data Science Consultant | Systems Limited

Apr 2022 - Dec 2024

- Deployed Arabic LLMs (LLAMA) on distributed GPU clusters using Azure ML. Optimized model serving and inference pipelines to significantly reduce latency and maximize hardware utilization for real-time analysis at scale.
- Designed and implemented time-series forecasting models on Azure Databricks for Estee Lauder, achieving a 15% increase in forecast accuracy through advanced feature engineering and hyperparameter tuning.
- Explored and researched a cloud-based clinical documentation tool designed to automate SOAP notes, medical coding, and administrative tasks for healthcare professionals, identifying integration paths for AI-driven workflow automation.
- Developed complex Power BI dashboards to surface critical insights from multi-billion-dollar datasets, supporting executive decision-making.

Senior Software Engineer, Data Analytics | Afiniti

Jan 2021 - Apr 2022

- Collaborated with research teams to deploy customer-agent pairing models into production environments, contributing to a 12% revenue increase for enterprise clients.

- Managed migration of extensive data pipelines to GreenPlum (PostgreSQL). Optimized SQL queries and ETL scripts, reducing data processing runtime by over 5 hours.
- Wrote reusable Python code for feature extraction and analysis, processing hundreds of attributes to improve model performance across batch and real-time workflows.

Lead Product Development | Stemers

Sep 2015 - Dec 2020

- Led the technical development of AI-based educational robotics kits, managing the full hardware-software lifecycle from ideation to mass production.
- Deployed educational technology solutions to over 40,000 users across Pakistan.

Focal Person for Pakistan (Volunteer) | Global Robot Cluster

Jan 2020 - Present

- Technical liaison facilitating knowledge exchange between Pakistani engineering stakeholders and the Global Robot Cluster network.

EDUCATION

MS in Computer Science | Institute of Business Administration, Karachi, Pakistan 2016 - 2019

BE in Electrical Engineering | National University of Sciences and Technology, Karachi, Pakistan 2011 - 2015

SKILLS

Languages & Frameworks: Python, SQL, PyTorch, TensorFlow, LangChain, LangGraph, LangSmith, FastAPI, REST APIs

AI & ML: Large Language Models (LLMs), Retrieval-Augmented Generation (RAG), Agentic AI, Agentic Workflows, Model Context Protocol (MCP), AI Development Kit (ADK), Embeddings, Vector Search, Computer Vision, NLP, Predictive Modeling, Time-Series Forecasting

Cloud & Infrastructure: Azure AI Foundry, Azure ML, Azure Databricks, Azure Cognitive Search, AWS, Docker, Model Serving, Inference Pipelines, Batch and Real-Time Inference, Scalable Infrastructure

MLOps & DevOps: MLOps, LLMOps, CI/CD Automation, Azure DevOps, Git, Observability and Monitoring, GenAI Pipeline Governance, Cost Optimization

Data & Visualization: Neo4j, Knowledge Graphs, Power BI

Domain: HealthTech, AdTech (Amazon Ads API), Financial Forecasting, Computer Vision

CERTIFICATIONS

- Azure AI Engineer Associate
- Azure Data Scientist Associate
- Azure AI Fundamentals