

M. Shaheer



Role: Sr. AI Engineer
Contact: info@botmer.io
Pricing per month: \$3500
Pricing per hour: \$30

PROFESSIONAL SUMMARY

AI Engineer with over 7 years of experience specializing in generative AI, NLP, and conversational AI solutions. Proven expertise in building and deploying LLM-powered systems, including retrieval-augmented generation (RAG) pipelines, multi-agent frameworks, and scalable AI architectures. Skilled in Python, LangChain, and cloud platforms like AWS, GCP, and Azure. Adept at working autonomously in startup environments and driving innovation through AI experimentation and education.

PROFESSIONAL EXPERIENCE

Senior Machine Learning Engineer

SOC Solutions pvt ltd – Karachi, Pakistan

May 2023 – Current

- Deployed and optimized LLM-based multi-agent systems using LangChain across healthcare and automotive sectors.

- Developed and maintained LangGraph, a custom framework for multi-tool integration and dynamic decision-making.

- Engineered conversational AI systems like Savant (healthcare) and Toyota Chatbot (automotive), achieving a 35% improvement in workflow efficiency.

- Built MLOps pipelines for real-time deployment and monitoring of Generative AI solutions.

Machine Learning Engineer

Senarios Private Limited – Lahore, Pakistan

Feb 2023 – Apr 2023

- Implemented time series forecasting models on AWS SageMaker, enhancing energy efficiency predictions by 25%.

- Deployed NLP-driven workflow automation for customer service, reducing response times significantly.

Research Assistant (AI & Machine Learning)

University of Essex – Colchester, UK
Oct 2023 – Jul 2024

Built LLM-based cybersecurity chatbot with LangChain, integrating RAG pipelines for accurate and context-aware responses.
Conducted experiments to advance natural language understanding and retrieval systems.

Junior Data Scientist

Unified Technology – Karachi, Pakistan
Aug 2018 – Jan 2023

Developed predictive time series models to optimize marketing campaign targeting, improving outcomes by 15%.
Designed seasonal and trend analysis pipelines to enhance operational decision-making.
Delivered robust analytics solutions to track and predict key business performance metrics.

ASR Engineer (Data Science & AI Projects)

Freelance – Karachi, Pakistan
Apr 2018 – Present

Delivered predictive analytics solutions, including neural network models for energy, healthcare, and real estate domains.
Developed multiple classification models, achieving high accuracy and scalable deployments.

AI Consultant

Freelance Projects – Remote
2019 – 2020

Designed and deployed conversational AI solutions for startups, leveraging OpenAI and Hugging Face models.
Conducted prompt engineering workshops and advised teams on integrating LLMs for scalable business solutions.
Improved customer engagement rates by deploying semantic search systems and FAQ bots tailored to specific industries.

EDUCATION

Master's in Data Science: University of Essex, Colchester, UK – Sep 2023
MSc in Computer Information Technology: NEDUET, Karachi, Pakistan – Sep 2022
Bachelor of Engineering in Computer information technology: NEDUET, Karachi, Pakistan – Sep 2019

PROJECTS

Energy Efficiency Prediction for Retrofitting

Developed multivariate time series models to forecast building performance, achieving a 40% reduction in environmental impact.

Integrated models into real-time systems via **AWS SageMaker**.

Savant Chatbot (Healthcare)

Built and deployed a **LangChain-based multi-agent chatbot**, enabling seamless healthcare professional support through advanced RAGpipelines.

Deployed the chatbot with advanced RAG pipelines to enhance retrieval and query response efficiency.

Toyota Chatbot (Automotive)

Developed a customer-facing chatbot for automotive support using LangChain, tailored to address user queries effectively

.Integrated advanced conversational workflows to streamline interactions and enhance user experience

Achieved a 35% improvement in workflow efficiency through optimized automation and intelligent response mechanisms

LangGraph Development

Built and deployed a custom **LangGraph framework** to handle multi-agent decision-making and tool integration.

Agent Performance Classification

Deployed **Multi label Classification model** for analysis the performance of the Agent on the basis of different parameters, in this we have using long former model and deployed it on sagemaker.

CERTIFICATIONS

Microsoft Azure AI900: Specialization in NLP and Image Processing

Microsoft Azure DP100: Advanced Data Science and Cloud-Based AI Deployment

GCP Tools and Deployment (Udemy)

SKILLS

Programming Languages: Python, SQL, Node.js, C/C++

NLP: Hugging Face Transformers, spaCy, Sentence Transformers

Machine Learning: Generative AI, LLMs, NLP (BERT, GPT), Time Series Analysis

AI System Architecture: Agent-Based Systems, Scalable AI Models, Cloud Deployments (AWS, GCP, Azure)

Development Tools: Git, Docker, CI/CD

Tools & Libraries: LangChain, LangGraph, TensorFlow, PyTorch, Docker, CI/CD

Cloud Platforms: AWS (SageMaker, EC2), GCP (Vertex AI), Azure (Azure ML)

Collaboration: Technical Communication, MLOps Pipelines, Multi-Tool Integration

Key Expertise: Semantic Search, Text Classification, Named Entity Recognition

ACHIEVEMENTS

Published research papers on **time series analysis** and **AI for sustainability** at international conferences.

Awarded 1st Prize in Web Development Competition (APTECH).