AMMAR WAQAR

+92 304 5448920 | Wah Cantt, Pakistan
• ammarwaqar.work@gmail.com • ammar-portfolio-ten.vercel.app
• linkedin.com/in/syedammarwaqar • github.com/dev-ammarrr

Summary

AI/ML Engineer with 2+ years of experience building and deploying end-to-end machine learning systems and LLM-powered applications. Proven expertise in Python-based ML pipelines, computer vision, and production-grade LLM integrations, including RAG systems, prompt engineering, and fine-tuning with LoRA/PEFT. Delivered AI solutions for national and international clients, including government-sector initiatives, and shipped automation workflows using LangChain, LangGraph, and n8n. Strong track record of taking models from experimentation to real-world deployment across NLP, computer vision, and applied AI products.

Skills

Programming	Python, C/C++, Java, JavaScript, SQL
Backend	FastAPI, Flask, REST APIs, Docker, Asynchronous Architectures
LLMs & RAG	Large Language Models (LLMs), Retrieval-Augmented Generation (RAG), LLM Fine-tuning (LoRA/PEFT), Prompt Engineering, AI Pipelines
Vector DB & Retrieval	Embedding Search, Semantic Retrieval, RAG Pipelines, Knowledge-Base Integration
Computer Vision	OpenCV, YOLO, Object Detection & Tracking, Image Processing, Video Analysis
Generative AI	GPT Models, Multimodal AI, LLM Application Development, Agentic Workflows
Frameworks	PyTorch, TensorFlow, Keras, Scikit-learn, Hugging Face Transformers, LangChain, LangGraph, OpenCV
RF Learning	Q-Learning, DQN, A3C, Multi-Agent RL
Automation	n8n, Supabase, Workflow Automation, AI-powered LLM Pipelines
Data Engineering	Data Preprocessing, Feature Engineering, Power BI Dashboards
Cloud & MLOps	Docker, Model Optimization, CUDA/cuDNN, Production Deployment
Frontend	HTML, CSS, JavaScript, REST API Integration
Soft Skills	Cross-Functional Collaboration, Client Communication, Technical Documentation, Stakeholder Communication, Iterative Prototyping

Work Experience

Senior Machine Learning Engineer
Sep 2026 – Present

Nysonian

- **Agentic AI Systems:** Built agentic AI systems that orchestrate customer-support workflows through tool use.
- **End-to-End ML Development:** Built models for varied use cases, handling the workflow from data collection and cleaning through training and iteration.
- **Customer Support Automation:** Developed and maintained customer-support agents and n8n workflows to automate support operations.
- **RAG and Prompt Reliability:** Improved retrieval and prompt reliability to help agents deliver accurate, context-aware responses.
- **Business-System Integrations:** Integrated tool calling with databases, Shopify, and ShipHero so agents can retrieve operational data and respond to customer inquiries.

Artificial Intelligence Engineer
2025 – Present

Intercraft *Islamabad, Pakistan*

- **Enterprise AI Delivery:** Developed AI-driven solutions for multiple national and international client projects across industries, covering problem framing, model development, and production integration.
- **Government-Sector AI Systems:** Contributed to AI system development for government-sector initiatives, including large-scale projects in Saudi Arabia.
- **Production ML Workflows:** Designed and integrated machine learning workflows for real-world applications, taking models from experimentation into production systems.
- **Applied LLM & Vision Systems:** Built production AI solutions across automation, computer vision, and LLMs.

Associate Software Engineer (LLM Trainer) 2025
Sirovista / Turing *Remote*

- **Human-in-the-Loop LLM Evaluation:** Evaluated LLM outputs as a human-in-the-loop judge to improve response accuracy and reliability across tasks.
- **Reasoning Quality Review:** Reviewed reasoning quality, factual correctness, and instruction adherence to guide iterative model improvements.
- **Prompt Evaluation Design:** Designed prompt evaluation scenarios to identify model weaknesses and reduce hallucinations.

AI & ML Intern Aug 2024
Arcelik R&D Center NUST *Islamabad, Pakistan*

- **Domain LLM Fine-tuning:** Fine-tuned pre-trained LLMs on domain-specific datasets, improving downstream task F1-scores by 15% across 3 R&D projects.
- **Prompt Templates & Model Variants:** Developed custom prompt templates and model variations, enhancing contextual accuracy in language generation tasks.
- **PEFT / LoRA Exploration:** Explored advanced fine-tuning strategies (LoRA, PEFT) under senior guidance, improving efficiency and real-world applicability.

Freelance AI/ML & Data Engineering Consultant Jan 2022 – Present
Fiverr & Upwork *Remote*

- **Client AI/ML Delivery:** Delivered 15+ AI/ML and data engineering projects for international clients, maintaining a 4.9/5 satisfaction rating.

Education

BS in Artificial Intelligence, University of Wah 2024
CGPA: 3.52

Key Projects

- **Autonomous AI Coding Agent for Full-Stack App Generation:** Built an autonomous multi-agent coding system (Python, LangGraph, LangChain, Docker, Claude) that generates complete full-stack applications (ERP, CRM, dashboards) from a single natural language prompt, with QA validation, Docker-based deployment, and automated debugging via conversational memory.
- **Dynamic Vision: Modular Real-Time Computer Vision System:** Developed a modular real-time computer vision system (Python, OpenCV) for CCTV and video streams, with detection modules for suspicious activity monitoring, construction worker safety compliance, and sports motion analysis.
- **Automated Book Generation System:** Designed an end-to-end n8n + GPT-4 pipeline (outline → chapters → compilation) that automated over 90% of the writing workflow, with Supabase for real-time storage and workflow tracking.
- **Automated Multi-Domain Customer Support System:**
  Built a voice-enabled multi-domain chatbot (ChatGPT-4, LangChain, Twilio, RAG) that handled 75% of customer queries autonomously, with domain-specific knowledge retrieval and fallback escalation to human agents.
- **Enhancing Robot Localization with Deep Learning:** Built a deep learning model (Python, TensorFlow, OpenCV) to solve the robot kidnapping problem from LiDAR and map inputs, achieving 93.5% accuracy within 20px and 97.4% within 25px on simulated datasets.
- **UFC Fight Analysis System:** Achieved 90% fighter detection accuracy with YOLO, automated logging of 500+ actions per match via MediaPipe and OpenPose, and reduced processing lag by 50% using CUDA/cuDNN.

Certifications

- IBM Machine Learning Professional Certificate
- AI For Everyone – Andrew Ng (Coursera)
- Supervised ML: Regression & Classification
- Unsupervised Machine Learning
- Deep Learning & Reinforcement Learning
- Exploratory Data Analysis for ML