SYED SAQIB

Full-Stack AI Engineer | AI Products End to End | Agentic AI · Agentic RAG · LLM Evals | Python · TypeScript · React · FastAPI | Claude Code · Cursor

syedsaqib313@gmail.com | linkedin.com/in/syed-saqib | +92-333-4141218 | Lahore, Pakistan

PROFESSIONAL SUMMARY

Full-stack AI engineer who ships AI products end to end, from client discovery to deployment: React and TypeScript front ends, FastAPI and SQL back ends, LLM agents with evals and guardrails, Docker and CI/CD on Azure and AWS. 6+ years across data science and predictive modeling (fraud incidents down ~18%, inventory holding costs down up to 22%), generative and agentic AI (multi-agent systems on Azure AI Foundry, agentic and corrective RAG) and full-stack AI products. Works AI-first every day in Claude Code and Cursor: founded a task-authoring kit of 68 tools and 17 agent skills (8,000+ commits) and built the operations platform for a 377-person AI data program.

SKILLS

Full-Stack AI Engineering: Python, TypeScript, SQL, React, Next.js, FastAPI, PostgreSQL, Redis / Celery, REST APIs
Generative & Agentic AI: LLMs, AI agents, multi-agent systems, LangGraph, MCP, agentic RAG, corrective RAG, structured outputs, function calling, evals (LLM-as-a-judge), Langfuse, guardrails, cost and latency optimization, agent memory
AI-Assisted Delivery: Claude Code (subagents, skills, hooks, plugins), Cursor (Bugbot), spec-driven development, context engineering, human-in-the-loop review of AI-generated code, GitHub Actions, Docker, CI/CD
Cloud: Azure (AI Foundry, Azure OpenAI, AI Search, Container Apps, Functions, Machine Learning), AWS (Lambda, Bedrock, S3, EC2), Vercel
Data Science & Predictive Modeling: Predictive modeling, classification, regression, time-series and demand forecasting, fraud and anomaly detection, statistical analysis, A/B testing, feature engineering, scikit-learn, Databricks
Client & Product: Product ownership, client discovery, pre-sales and solution consulting, team leadership (up to 377)

PROFESSIONAL EXPERIENCE

Turing | Remote | Contract — Mar 2024 – Present
Consultant – AI Engineer & Technical Lead (AI Tooling & Automation)

- **AI-first engineering:** Builds daily in Claude Code and Cursor with human-in-the-loop review of AI-generated code; founded the team's task-authoring kit (8,000+ commits; 68 Python tools, 17 agent skills, 2 safety hooks) and first-authored 48 of its 82 tool scripts.
- **Operations platform:** As Team Lead of a 377-person AI data program, built 131 Python automation scripts (~34K lines) for onboarding, delivery and dashboards, plus a FastAPI, SQLAlchemy and React/TypeScript review app.
- **Multi-agent coding factory:** Ran a coordinator plus 5 autonomous Claude Code agents (3,500+ commits) that delivered 171 of the team's 439 RL tasks (~39%) for an enterprise data & AI platform's autonomous data agent.
- **LLM-as-a-judge QA tool:** Built a ~1,800-line Python judge on an LLM API for duplicate detection, live catalog collision checks and taxonomy review, packaged for other authors; its verdicts cover 92 of the 103 first-phase tasks.
- **CI quality gates:** Shipped 15+ automated gates (contamination, trace integrity, rubric alignment, readiness judge) and co-led a data-estate expansion with zero breakages across 389 regression-tested tasks and 990 graded runs.
- **Earlier programs:** RLHF and evaluation data for a frontier AI lab (1,500+ code-verified questions, 2,300+ rated A/B comparisons); as Project Lead, 1,260 Kaggle competitions in MLE-bench format in 12 days; 13 auto-graded ML benchmark problems on data of up to 33.8M rows.

MindHYVE.ai | Remote | Contract — May 2025 – Jan 2026
Agentic AI Engineer

- Designed and deployed end-to-end agentic AI systems and Azure-native inference pipelines on Azure AI Foundry (Azure ML, Azure Functions, Azure AI Services) with LangGraph, CrewAI and Semantic Kernel.
- Automated LLM deployment and prompt optimization with Azure PromptFlow (LLMOps), reducing errant outputs by an estimated ~37% in production tests; traced and evaluated pipelines with Langfuse and LangSmith.
- Built agentic RAG over Microsoft 365 content with corrective / self-RAG loops that re-retrieve when an answer is not grounded in company knowledge.

Royal Cyber Inc. | Lahore, Pakistan Dec 2022 – May 2024
Senior Data Scientist

- Led pre-sales client calls and solution proposals across North America, Canada and the Middle East: discovery of client challenges, tailored AI solution design and negotiation.
- Fine-tuned LLMs (GPT, LLaMA, Falcon, Mistral) with LoRA and PEFT and put them to work in enterprise document search (FAISS, Pinecone), legal and financial summarization, and customer-support chatbots.
- Led a team of junior engineers building Generative AI POCs for real-time applications.
- Built fraud and anomaly detection and demand-forecasting models for retail and financial clients (fraud incidents down ~18%, inventory holding costs down up to 22%) and supply-chain forecasting for a US beverage manufacturer (throughput and efficiency up ~15%), deployed through CI/CD on Azure and Databricks.

Fiverr & Upwork | Remote | Freelance Feb 2020 – Dec 2023
Python Developer & Machine Learning Specialist (Fiverr); Data Scientist (Upwork)

- Client delivery from scoping to hand-off: 120+ Fiverr projects as a 5-star Level 2 seller (Feb 2020 – Nov 2022) and 25+ Upwork projects with a 100% success rate (May 2022 – Dec 2023), across Python, ML, NLP, OCR and Power BI.

xiQ Inc. | Lahore, Pakistan Jul 2020 – Dec 2022
Data Scientist

- Built hybrid search (Elasticsearch + vector search) with Learning-to-Rank re-ranking for the xiQ SalesApp and Workbench, improving search relevance metrics by ~25%.

Earlier: front-end and back-end web development internships at 360Dynamics and Powersoft19, Lahore (2017–18)

PROJECTS

Governed Analytics Agent ("Ask Your KPIs"): Full-Stack AI Product | Independent project (2026)

- Product owner and sole builder, from idea and UX to deployment: Next.js (TypeScript) chat UI with streaming answers and charts, FastAPI backend, Clerk auth, deployed on Vercel and Azure Container Apps.
- Function calling into typed metric APIs over a dbt semantic layer, so the model never writes SQL; row-level security, PII masking, Redis session memory, a forecasting tool, Langfuse traces and tool-call evals.

Agentic RAG Service with MCP Tools and Eval-Gated CI | Independent project (2026)

- FastAPI + LangGraph agent calling tools through MCP servers; hybrid search with reranking, a corrective-RAG grader loop and Celery ingestion workers; LLM-as-a-judge eval gate in GitHub Actions; semantic caching and model routing for cost and latency; deployed on Azure Container Apps (Azure OpenAI) and Amazon Bedrock.

Terminal-Bench-Style Task Pack and Agent Harness | Independent project (2026)

- Dockerized coding tasks with pytest checkers and an Inspect harness that runs Claude Code and a Codex-style CLI agent, logs trajectories, classifies failures and tests every grader against perfect and broken solutions.

EDUCATION

MS Data Science | National University of Computer and Emerging Sciences (FAST NUCES) Aug 2022 – May 2024
CGPA 3.7 | Silver Medalist (2024) for academic performance and excellence in research

BS Computer Science | FAST NUCES | CGPA 3.31 | Dean's List Aug 2016 – Jul 2020

CERTIFICATIONS

- Coursera Specializations: Machine Learning; Deep Learning; Machine Learning Engineering for Production (MLOps); Natural Language Processing
- Generative AI for Large Language Models (LLMs); Building Agentic RAG with LlamaIndex; Functions, Tools and Agents with LangChain; LangChain for LLM Application Development; Building Multimodal Search and RAG; LLMOps

AWARDS

- Star Performer of the Year, Royal Cyber (2023) | Top Performer, xiQ (2021–22) | ICPC Asia Regional, 6th (2019)