Muhammad Faras Siddiqui

LinkedIn: linkedin.com/muhammad-faras-siddiqui
GitHub: github.com/siddiquifaras
Mobile: +971 56 8815205 / +92 335 9048481
Email: siddiquifaras@gmail.com
GitLab: gitlab.com/sidfaras

Summary

AI Engineer with hands-on experience in infrastructure, backend systems, applied ML, and deployment-heavy environments. Built and scaled production-grade AI systems, designed infra architecture from scratch. Carry a strong bias for pragmatic execution and crafting end-to-end solutions.

Education

- **National University of Science and Technology (NUST)** — Islamabad, Pakistan
  Bachelor of Engineering in Software Engineering — *Nov. 2021 – Jun 2025*

Experience

- **Founding Engineer** — June 2025 – Present
  CVision — *Islamabad, Pakistan*
  - Built and deployed distributed systems powering client-facing AI applications.
  - Drove $70K+ in client revenue over first 6 months through infrastructure and systems delivery.
  - Owned CI/CD pipelines, devops stack, and backend infra across products and internal tooling.
  - Delivered truth-grounded retrieval and knowledge pipelines for operational automation.
  - Led R&D initiatives across GNNs, knowledge graphs, and efficient architecture design.
  - Built low-latency, spike-driven inference libraries enabling compression ratios of 1000% for deep networks.
- **Deep Learning Research Intern** — June 2024 – Sept 2024
  Optical Networks and Technologies Lab — *Islamabad, Pakistan*
  - Designed and trained models for optical signal classification and anomaly detection.
  - Used GANs and transformers for signal robustness and sequence-level interpretability.

Key Projects

- **Neuromorphic Compression and Tooling Stack**:
  - Developed 3 standalone Python libraries for spike-based neural computation.
  - Researched and proposed compression via custom serialization formats (ex: from 200KB to 200B for 10-layer networks) .
  - Built primitive simulation engine, IR visualizer, and binary encoder for efficient inference.
- **CVision - Virtual Try-On Engine**:
  - Engineered geometry-based try-on prototype with real-time CV + ML integration.
  - Handled backend service orchestration and real-time data pipelines.
- **Bizmila - AI Co-Founder Platform**:
  - Led backend (Node.js, TypeScript), Python APIs (FastAPI), and Dockerized infra.
  - Delivered multi-agent workflows turning startup ideas into investor-ready deliverables.
  - Integrated multiple LLMs (OpenAI, Claude, Cohere) with custom evaluation + fallback logic.

Skills

- **Languages:** Python, C++, JavaScript, TypeScript
- **Systems/Infra:** Docker, CI/CD, GitHub Actions, Linux, Kubernetes (basic), self-hosted ops
- **Backend:** Node.js, Express, FastAPI, Flask, REST APIs, MongoDB, MySQL
- **AI/Tooling:** PyTorch, SNNs, GNNs, TensorFlow, LangChain, knowledge graphs, RAG
- **Other:** Git, GCP, Azure, testing pipelines, prompt frameworks