Zubair Atha

+1 646-392-5601 | zubair.atha09@gmail.com | linkedin.com/in/zubair-atha | github.com/zubairatha

Education

Columbia University — **New York, NY**
MS in Data Science — *Aug 2024 – Dec 2025*
- **Courses**: Machine Learning; Algorithms; Probability; Statistical Modeling; Deep Learning; NLP; Finance; LLM based GenAI
- **Graduate Teaching Assistant**: Python for MBAs, Introduction to Databases, Data Analytics
- Data Science Institute Scholar [Spring 2025]

Vellore Institute of Technology — **Vellore, India**
BTech in Computer Science and Engineering — *Sep 2020 – Jun 2024*
- **Courses**: Data Structures; Operating Systems; Databases; Cloud Computing; Artificial Intelligence; Distributed Systems
- Raman Research Award Recipient [Fall 2023]

Technical Skills

GenAI & LLMs: Agentic AI, Multi-Agent Systems, RAG, MCP, A2A, Search and Retrieval Systems, LLM Evaluation, RLHF
ML & Deep Learning: Transformers, Fine-Tuning (LoRA, PEFT), XGBoost, CNNs, Regression, Classification, Clustering, PCA
Frameworks & Tools: PyTorch, LangGraph, LangChain, Pydantic, Numpy, Hugging Face, Qdrant, FAISS, Scikit-learn, FastAPI
Engineering: Python, SQL, C++, PostgreSQL, Javascript, Docker, REST APIs, CI/CD, Distributed Systems, GCP, AWS, Spark

Experience

Ford Motor Company — **Dearborn, MI**
Data Scientist — *Jan 2026 – Present*
- Co-leading agentic AI system design and development for a **0-to-1 vehicle budget planning product** under a **$50M+ AI initiative**; driving research in **Google ADK harnesses, search, prompt optimization (GEPA), and LLM evaluation**.
- Architected and deployed a **GCP Cloud Run orchestrator-worker pipeline** processing **25K events/250 GB daily in 30 min** with sharding, parallelization, concurrency, idempotent **BigQuery** merges and **Docker/Cloud Build**.
- Migrated **3 production pipelines** from AWS S3 to internal APIs (**12K logs/120 GB daily**), cutting runtime **86% (22h to 3h)** with 3-tier failover, circuit breakers, retries and connection pooling; eliminated manual recovery.

Ford Motor Company — **Dearborn, MI**
Data Science Intern — *May 2025 – Aug 2025*
- Engineered a **Gemini 2.5-powered text-to-code platform** translating natural-language queries into **sandboxed, executable regex code** to analyze **10K+ ECU logs**, reducing engineer-analyst workflow time by **90%**.
- Developed a **LangGraph multi-agent root cause analysis system** with **3-layer hybrid RAG**, JSON parsing, **Qdrant/HNSW** and an **LLM eval framework** (golden set, rule checks, LLM-as-a-Judge), achieving **85% accuracy**.
- Cut log-query latency **99% (10s to 0.1s)** with **Redis caching and multithreading**; contributed to a **patent filing** for the end-to-end GenAI codegen, hybrid retrieval and ECU-log analytics framework.

Columbia University Irving Medical Center — **New York, NY**
Graduate Researcher — *Jan 2025 – May 2025*
- Built a **computational neuroscience pipeline** with **DeepLabCut**, time-series features and **KMeans/HDBSCAN**, identifying **16 behavioral clusters** in mice covering **80%** of frames across 20+ videos.

56 Secure — **Bengaluru, India**
Machine Learning Engineer Intern — *Feb 2024 – May 2024*
- Trained **YOLOv5s** on **500K+ images** using **PyTorch**, deploying optimized ONNX models across **500+ AI cameras**.
- Engineered **active learning and knowledge distillation pipelines**, resolving **30+ production detection issues**.
- Implemented a **model-evaluation slack bot** using statistical modeling flagging **150+ field incidents**.
- Automated **Python/SQL/AWS S3/Spark** data annotation pipelines, accelerating training-data preparation by **3x**.

Vellore Institute of Technology — **Vellore, India**
Undergraduate Researcher — *Jan 2023 – Oct 2023*
- First-authored an **IEEE Access** paper on **SSBTCNet**, a semi-supervised brain tumor MRI classifier using autoencoders for feature learning and fuzzy logic image enhancement, achieving **97.3% accuracy with 45% unlabeled data**.

Projects

CalBridge - Intelligent Calendar Assistant | *Python, Qwen2.5, Ollama, FastAPI, PyObjC, SQLite*
- Orchestrated an **8-stage agentic calendar assistant** using **Qwen2.5-14B and Ollama** to interpret scheduling requests, resolve constraints, and schedule events through **FastAPI** on my macOS Calendar.

AI Email Automation Agent | *Python, Ollama, GPT-4, RAG, Graph RAG*
- Developed a **Chrome extension** that drafts personalized emails across various email platforms using **local Ollama LLMs**, **RAG/Graph RAG** over past emails and memories, and GPT-4-generated FAQs, achieving **∼2s draft latency**.