

MUHAMMAD MUJTABA AMJAD

AI/ML Engineer · Computer Vision · Big Data Systems · Open-Source Contributor

+92 333 130 6771 · mmujtaba0085@gmail.com · mujtaba0085.opior.com · github.com/mmujtaba0085 · [LinkedIn](#) · Islamabad, PK

PROFILE

CS graduate from FAST NUCES with a habit of building things that actually ship. Merged PR #8591 into NetworkX (16.8k stars) — a butterfly counting algorithm running **up to 16× faster** than the existing implementation — after working through design review with three core maintainers. Beyond that: a statechart hosting platform backed by a formal versioning proof, a containerised agentic AI platform, a 14-service crime analytics pipeline on Spark and Storm, and an 8-layer AI construction monitoring system with NL→SQL. The common thread is an obsession with getting the details right, whether that is cache contention in a distributed log or bcrypt cost factors in an auth flow.

EDUCATION

FAST – National University of Computer & Emerging Sciences, Islamabad

2022 – 2026

BS Computer Science · Major: Data Science, Deep Learning, Machine Learning

OPEN-SOURCE CONTRIBUTIONS

NetworkX — Merged PR #8591: butterflies() for the bipartite module

Mar 2026

Python, C++, MPI, OpenMP

- Achieved **up to 16× speedup** over NetworkX's existing implementation on sparse and skewed bipartite graphs, benchmarked across KDD 2018 and VLDB 2023 algorithm variants.
- Proposed, implemented, and merged **butterflies()** into NetworkX (16.8k stars); built the underlying C++/MPI/OpenMP parallel counter from scratch (PARBUTTERFLY-inspired) before contributing the Python API to the bipartite module.

PROJECTS

StateKeep — Versioned Statechart Hosting Platform

2025 – Present

TypeScript, XState v5, Node.js, C, Formal Methods

- First statechart hosting platform to solve live actor migration: workflow machines can be updated while actors are running, each routed deterministically to the correct state in the new version — no replay, no migration scripts, no side-effect re-firing. Resolves an open problem the XState maintainer confirmed had no formal solution since 2020.
- Designed Anchor-Point Versioning (APV): routes in-flight actors by tracing stable execution checkpoints rather than current state alone — validated at 10^6 actors with 84.86% memory reduction and 90.94% routing speedup over event-sourcing replay; C migration engine at $p50\ 1.26\ \mu s/actor$, $\sim 97k$ actors/sec throughput.

ConstructionSightAI — Final Year Project

2024 – 2026

Python, YOLOv8, OpenCV, FastAPI, React, PostgreSQL, JWT/OAuth, SSE

- GPU/CUDA batched inference over live RTSP streams: detection, PPE compliance, multi-object tracking, cross-camera re-identification, and activity state machines feeding five intelligence engines (PPE, Workforce Analytics, Equipment, Activity, Risk) via SSE and PostgreSQL persistence.
- FastAPI backend with OAuth/JWT, role-based access control (five roles); React frontend with 8 dashboards and a Smart Query Assistant translating natural language to validated SQL with AI-explained answers and charts.

Echo-Persona — Digital Twin AI Platform

Dec 2025

Python, FastAPI, Streamlit, Whisper, ChromaDB, Groq API, Docker

- Containerised persona platform with RAG (ChromaDB + Groq llama-3.3-70b), offline Whisper STT (English & Urdu), AI voice cloning, and two agentic models; multi-stage Docker build reduced image from ~ 8.5 GB to ~ 3.3 GB; 20+ REST endpoints, WebSocket transcription, multi-format voice output.

Real-Time Crime Analytics & Intelligent Alert System

May 2026

Python, Spark, Kafka, Storm, PySpark MLlib, PostgreSQL, MongoDB, Docker

- Lambda Architecture over 50k+ Chicago crime records: Spark batch layer with K-Means hotspot detection and a Kafka + Storm 5-bolt pipeline generating 6 530+ severity-graded alerts across 22 districts; dual serving in PostgreSQL + MongoDB; 14 Dockerised services; 9-page Streamlit dashboard with 30s auto-refresh.

Knack — Freelancer Marketplace

2025

Node.js, Express, MySQL, EJS, Multer, Nodemailer

- Three role-separated dashboards (Client, Freelancer, Admin) with proposals, milestone-based contracts, in-contract messaging, and a dispute queue; production auth: bcryptjs (cost 12), hashed reset tokens, rate limiting, express-validator sanitisation; 18-table schema.

TECHNICAL SKILLS

Languages

Python, JavaScript (Node.js), TypeScript, C++, Java, SQL, PL/SQL

ML / CV

PyTorch, TensorFlow/Keras, OpenCV, YOLOv8, Detectron2, Whisper, scikit-learn, NumPy, Pandas

Big Data

Apache Spark, PySpark MLlib, Kafka, Apache Storm, Lambda Architecture, Zookeeper

Backend / Web

FastAPI, Flask, Express.js, React, XState, REST APIs, RAG, ETL, Groq API

Databases

PostgreSQL, MySQL, MongoDB, SQLite, ChromaDB

DevOps / Tools

Docker, Docker Compose, AWS (EC2, S3), Git, MPI, OpenMP, METIS, Streamlit, n8n