

AMEER HAMZA NASIR

Data Scientist · Python / ML / Statistical Modeling

hamza6700@gmail.com | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

SUMMARY

Data Scientist and Computer Science student with hands-on experience building and deploying machine learning systems across fraud detection, computer vision, and predictive modeling. Achieved **F1 = 0.91 and AUC-ROC = 0.98** on a highly imbalanced credit-card fraud dataset and reduced simulated traffic intersection wait time by **36%** through real-time vehicle detection and adaptive scheduling. Experienced across **Python, SQL, Scikit-learn, PyTorch, XGBoost, FastAPI, Docker, AWS, and CI/CD**.

TECHNICAL SKILLS

Languages: Python, C++, JavaScript, SQL, Bash

ML & Data Science: Scikit-learn, PyTorch, Pandas, NumPy, EDA, Feature Engineering, Statistical Analysis, Model Evaluation (R^2 , MAE, RMSE, F1, AUC-ROC)

Deep Learning & CV: Neural Networks, CNN, OpenCV, YOLOv5, Object Detection, Image Processing

MLOps & Deployment: FastAPI, Flask, Streamlit, Docker, Kubernetes, GitHub Actions (CI/CD), REST APIs, Cloud Deployment (AWS EC2)

Data & Tools: SQL, SQLite, Git, Linux, Shell Scripting, Jupyter, Pytest

KEY PROJECTS

Credit Card Fraud Detection System *Python · XGBoost · FastAPI · Docker · AWS*

- Achieved $F1 = 0.91$ / $AUC-ROC = 0.98$ by benchmarking Logistic Regression, Random Forest, and XGBoost on a highly imbalanced dataset (492 fraud / 284,807 transactions), using SMOTE oversampling, stratified k-fold cross-validation, and threshold tuning.
- Shipped a production-ready inference endpoint by deploying the model via FastAPI, containerized with Docker on AWS EC2, with Bash-scripted reproducible deployment.
- Identified the strongest fraud signals by conducting full EDA, transaction-distribution analysis, correlation heatmaps, and PCA visualization of feature clusters.

Intelligent Traffic Signal Optimizer *Python · PyTorch · OpenCV · YOLOv5*

- Cut average intersection wait time by 36% by designing a real-time adaptive traffic signal system with a Greedy scheduling algorithm that dynamically allocates green time based on live vehicle density, replacing Pakistan's fixed 60–120s timers.
- Enabled real-time per-lane vehicle counts by training a YOLOv5 (PyTorch) detection model on Pakistani road, processed via OpenCV frame capture.
- Reduced congestion at high-density lanes by implementing a priority-queue (min-heap) scheduler in Python to rank and serve lanes by congestion score.
- Surfaced signal performance in real time by building a Python simulation dashboard tracking state transitions, queue lengths, and throughput across a 4-way intersection.

Student Exam Score Predictor *Python · Scikit-learn · FastAPI · Docker · CI/CD*

- Predicted exam outcomes at $R^2 = 0.68$ / $MAE = 2.03$ by building a full ETL pipeline (ordinal + binary encoding, feature-selection ranking, MinMax scaling) over 6,607 student records and training a Random Forest Regressor.
- Guaranteed pipeline reliability by writing pytest unit tests for ETL and model outputs, integrated into GitHub Actions CI, and deployed via Streamlit and a Dockerized FastAPI endpoint.

Claudecode-free — Open-Source *Node.js · Docker · GitHub Actions*

- Eliminated Anthropic API billing for users by architecting an isolated Docker sandbox that routes Claude Code API calls through a custom Node.js reverse proxy (proxy.js) to OpenRouter free-tier LLMs.
- Earned organic GitHub stars from developers worldwide with zero promotion by engineering real-time Anthropic↔OpenRouter format translation with request deduplication and SSE streaming, and automating Docker builds via GitHub Actions CI/CD.

EDUCATION

Government College University, Faisalabad — B.S. Computer Science

May 2029

Relevant coursework: C++ (OOPS, DSA,) OS, Calculus, ICT, Relational Databases & SQL

LANGUAGES

English — Professional, Urdu — Native, Punjabi — Native, Deutsch — beginner