YUSUF MUHAMMAD MUSA

AI Systems Engineer | Operations Research & Optimization | Data & ML Engineering

Bauchi, Nigeria · yusuf2000mm@gmail.com · +234 814 400 3669

yusuufmm.is-a.dev · github.com/yusuuf-mm · linkedin.com/in/yusuufmm

PROFESSIONAL SUMMARY

AI Systems Engineer with a B.Tech in Operations Research, focused on the substrate that sits underneath a model: the data pipelines that feed it, the Mathematical Optimization layer (MILP, LP via PuLP, Pyomo, Google OR-Tools) that turns its output into a decision, and the agent orchestration layer (LangGraph, NVIDIA NIM) that coordinates the whole system. Built four end-to-end Decision Intelligence platforms spanning energy, logistics, and behavioral scheduling — each pairing a deterministic solver with an AI reasoning layer, backed by Kafka/Spark/Airflow/dbt orchestration and FastAPI/Docker/Terraform infrastructure. Seeking Junior-to-Mid AI Systems / ML / Data Engineering roles where the model is one component of a larger production system, not the whole product.

TECHNICAL SKILLS

AI Systems & LLMs: RAG, LangGraph, LangChain, Agent Orchestration, NVIDIA NIM, Gemini API, OpenAI API, Vector Databases (Qdrant), fastembed, MLflow

Operations Research: MILP, LP, Mathematical Optimization, Monte Carlo Simulation, Solvers (PuLP, Pyomo, Google OR-Tools, CBC, HiGHS, GLOP), DEA, Markov Chains

Data & Backend Engineering: Apache Kafka, Spark, Airflow, dbt, DuckDB/MotherDuck, FastAPI, Flask, REST APIs, Pydantic, SQLAlchemy (async), Celery, OpenAPI/Contract-First Design

MLOps & Infrastructure: Docker, Terraform, AWS (S3, EC2, Lambda, DynamoDB, API Gateway), GitHub Actions (CI/CD), Nginx, PostgreSQL, Redis, BigQuery

Languages & Tools: Python, SQL, TypeScript/JavaScript, C, Bash, VBA, Streamlit, Plotly, Evidence.dev, Git/GitHub

SELECTED PROJECTS

HERMES — Multi-Agent Logistics Decision Intelligence Platform

Python · Google OR-Tools · LangGraph · NVIDIA NIM · DuckDB

- Cut ~70% of LLM agent calls during normal operations by building a conditional-edge LangGraph pipeline (5 agents: Monitor, Classify, SLA-Risk, Reroute, Dispatch) on NVIDIA NIM that skips reasoning steps when no disruption is detected.
- Solved a 26-node, 5-vehicle routing problem to 171.96 km with zero constraint violations via a Google OR-Tools CVRPTW solver under 7 hard constraints (capacity, time windows, shift limits, depot routing), validated by a 7-test suite.
- Eliminated duplicate event processing across unlimited pipeline re-runs by designing an idempotent DuckDB anti-join pattern (LEFT JOIN WHERE NULL) for the event-driven orchestration layer.
- Shipped a 6-page live operations dashboard by syncing DuckDB to MotherDuck and auto-building/deploying an Evidence.dev site to GitHub Pages on every push via GitHub Actions.

NETI-HyOptima — Net-Zero Energy Transition Decision Intelligence Platform

Python · Pyomo · HiGHS/CBC · Monte Carlo Simulation

- Built a MILP engine in Pyomo (solved via HiGHS/CBC) that sizes and dispatches hybrid solar-gas-battery systems to minimize total cost (CapEx, fuel, grid, penalties) under power-balance, ramp-rate, and reliability constraints.
- Fixed a sizing bug that let the solver assign unlimited charge/discharge power to an undersized battery by re-deriving C-rate constraints to bind to the installed-capacity decision variable, then applying Big-M linearization to remove bilinear terms.
- Quantified investment risk across 4 uncertainty dimensions (fuel cost, carbon price, demand growth, solar capacity factor) via a parallelized Monte Carlo layer computing VaR, CVaR, and technology-adoption probabilities.
- Validated against 6 Nigeria-specific location profiles and 4 scenario archetypes aligned to Nigeria's Energy Transition Plan, with ~60 unit tests covering optimization and simulation modules.

EcoGrid-Agent — AI-Orchestrated Grid Optimization System

Python · FastAPI · Celery · Gemini SDK · Qdrant · Google OR-Tools

- Converted natural-language grid requests into optimal 24-hour battery charge/discharge schedules by orchestrating a FastAPI → Celery → Gemini SDK pipeline that calls a deterministic solver as the final step.
- Removed solver-constraint values from source code by building a RAG pipeline — fastembed embeddings over Qdrant — that retrieves regulatory safety buffers from policy documents at runtime.
- Implemented a 24-hour Linear Program (Google OR-Tools GLOP, 5 constraint families: battery physics, capacity, rate limits, solar curtailment) supporting 3 selectable objectives: maximize profit, minimize cost, minimize carbon.
- Produced a fully audited decision trail (every tool call logged with input, output, and duration) via an async FastAPI/Celery model polled through Redis.

HabitOS — AI-Driven Behavioral Decision Support Platform

FastAPI · PuLP/CBC · PostgreSQL · React · TypeScript

- Returned optimal daily schedules in under 500 ms for 10-15 candidate behaviors by formulating a MILP (PuLP + CBC) that maximizes a weighted "life value" objective under time-budget, energy-budget, time-slot, and frequency constraints.

- Kept a FastAPI/Pydantic backend and React/TypeScript frontend in permanent type-sync through a contract-first architecture around a 991-line OpenAPI specification.
- Automated test, build, and deployment via GitHub Actions CI/CD running unit and integration tests against ephemeral PostgreSQL/Redis containers, with Render deployment gated on health checks.
- Implemented JWT access/refresh auth, async SQLAlchemy 2.0 with Alembic migrations, and Redis-backed caching for the production data layer.

ADDITIONAL PROJECTS

- **Real-Time Energy Data Pipeline** — Achieved ~3% transmission loss across a simulated 5-zone grid via a two-stage PuLP pipeline (fairness LP + transportation-problem LP). Streamed via Kafka/PySpark into PostgreSQL and S3, orchestrated by a 10-minute Airflow DAG (freshness check → dbt run → optimizer), with AWS infrastructure via Terraform and results in a Streamlit/Plotly dashboard.
- **Titanic Survival & Lifeboat Optimization Engine** — Two-stage decision system: XGBoost classifier (85% accuracy, 0.82 F1) serves calibrated survival probabilities via FastAPI to a PuLP MILP that allocates lifeboat seats under capacity, demographic-priority, and family-cohesion constraints. MLflow model registry with S3 artifact storage and DynamoDB-backed prediction logging and rate limiting.

PROFESSIONAL EXPERIENCE

ALX Software Engineering Program Oct 2024

Software Engineer & Portfolio Project Reviewer *Remote*

- Raised code quality across a global cohort by reviewing backend, API integration, and database-design submissions for Git/GitHub version-control and coding-standard adherence, with written technical feedback on each.
- Improved junior developers' command of production patterns through mentoring on backend development, ELT/ETL pipeline design, REST API design, and PostgreSQL data modeling inside agile sprint-based workflows.

Gombe State Investment & Property Development Co. / GDSS Feb 2024 – Feb 2025

Operations Analyst & Educator (NYSC) *Gombe, Nigeria*

- Gave decision-makers a clearer view of real-estate investment trends by querying property data with SQL and building Tableau dashboards for financial-planning and resource-allocation discussions.
- Improved reliability of strategic-planning inputs by applying data governance and data-quality checks to underlying datasets.

Ministry of Budget and Economic Planning, Bauchi State Dec 2022 – Jun 2023

Operations Research Intern *Bauchi, Nigeria*

- Cut manual data-processing effort by ~80% and analysis time by 60% by building a Decision Support System in Excel/VBA for the state's planning team.
- Identified 8 efficient and 12 underperforming LGAs with redistribution recommendations by applying DEA, Linear Programming, and Markov Chain models to teacher-allocation and 5-year enrollment data.

National Youth Service Corps (NYSC), Amada Camp Feb 2024

Operations Research Analyst *Gombe, Nigeria*

- Drove 15% reduction in complaints, 20% faster medical response, and 25% efficiency improvement by designing and running a 408-respondent operational survey (90.67% response rate, Excel/SPSS/R) and presenting recommendations to camp leadership.

EDUCATION

Modibbo Adama University, Yola Graduated Jan 2024

B.Tech (Hons), Operations Research *Yola, Nigeria*

- Relevant coursework: Linear & Nonlinear Programming, MILP, Dynamic Programming, Goal Programming, Network Flow & Graph Theory, Simulation Techniques, Decision Theory, Probability & Regression Analysis, Artificial Intelligence, Management Information Systems.
- Coursework directly underpins the LP/MILP solvers, DEA, Markov models, and Monte Carlo simulations used across all featured projects.

CERTIFICATIONS & TRAINING

- **Data Engineering Zoomcamp** — DataTalks.Club: 9-week project-based program covering Docker, Terraform, Airflow/Kestra/Prefect, BigQuery, dbt, Spark, and Kafka, capped by an end-to-end capstone pipeline.
- **Data Engineering (DeepTech)** — Data Science Network (DSN) / 3MTT: Intensive program focused on architecting production-grade data substrates; implemented batch and streaming pipelines utilizing Apache Airflow and Apache Kafka for real-time telemetry, with enterprise ETL orchestration via Azure Data Factory.
- **Machine Learning Engineering Zoomcamp** — DataTalks.Club: End-to-end ML pipeline development, evaluation, and deployment with FastAPI, Docker, Kubernetes, and AWS Lambda.
- **AI Development Tools Zoomcamp** — DataTalks.Club: AI-assisted development workflows using GitHub Copilot, Claude, and ChatGPT, including agent-based automation and CI/CD integration.
- **LLM Zoomcamp** — DataTalks.Club (in progress): Large language models, RAG systems, fine-tuning, and deployment.
- **Cloud Engineering (AWS)** — SchullTech: EC2, S3, Lambda, and API Gateway for scalable, serverless, data-driven applications.
- **ALX Software Engineering Program**: Backend development, API design, and scalable system architecture.