

SABINA BIMBI

AI/ML Engineer · Data Scientist · Full-Stack Developer

sbimbi@mtu.edu · linkedin.com/in/sabina-bimbi · github.com/SabinaLucy · huggingface.co/SabinaLucy

PROFESSIONAL SUMMARY

Analytical Data Science Master's student specialising in Machine Learning, Deep Learning, and Agentic AI. Proven track record designing and deploying production-grade ML systems, from a live 870K-headline financial intelligence platform to unsupervised infrastructure anomaly detection. Expert in Python, PyTorch, and TensorFlow, with full-stack engineering experience across FastAPI, React, Docker, and AWS. Adept at working through ambiguity to deliver scalable, end-to-end technical solutions that cross the gap between research and production. Open to AI/ML Engineering, Data Science, Software Engineering, Data Analytics, and Business Intelligence roles.

EDUCATION

Master of Science, Data Science

Dec 2025 – May 2027

Michigan Technological University · Houghton, MI

Specialisation: AI/ML · GPA: 3.75/4.0 · Deep Learning · NLP · Advanced Data Mining

Bachelor of Science, Information Technology

2020 – May 2024

Richfield Graduate Institute of Technology · Johannesburg, South Africa

Specialisation: Software Engineering

TECHNICAL SKILLS

Languages: Python · JavaScript · TypeScript · SQL · R · Java · C# · C++ · PHP · HTML/CSS

ML / AI: PyTorch · TensorFlow/Keras · XGBoost · LightGBM · Scikit-learn · Conformal Prediction · HMM · YOLOv8 · ViT

Generative AI & APIs: Claude API (Anthropic) · OpenAI GPT-4o · Gemini API · LangChain/LangGraph · ChromaDB · MCP Servers · FRED API · NewsAPI

Engineering & DevOps: FastAPI · React 18 · Node.js · Express · Docker · AWS (EC2, ECR, DynamoDB) · HuggingFace Spaces · GitHub Actions

Data & Enterprise: Tableau · Power BI · Oracle DB · PostgreSQL · MySQL · CSGold · TouchNet · Confluence · Seaborn · Matplotlib · Excel

Domains: Healthcare AI · Financial ML · Agentic AI · Time Series · NLP · Computer Vision · Infrastructure Monitoring

PROFESSIONAL EXPERIENCE

IT Operations & Documentation Specialist

Dec 2025 – Present

Michigan Technological University · Houghton, MI

- Monitor and troubleshoot campus-scale financial infrastructure across **TouchNet, DSE, and Genius** transaction platforms, ensuring high system availability and secure processing.
- Execute exploratory data audits and query records within enterprise **Oracle and CSGold** database environments under strict PII confidentiality protocols.
- Build custom **Python** utility scripts to parse raw server logs and automate routine data validation tasks, reducing manual review overhead.
- Author technical documentation, system blueprints, and manual override procedures in Confluence to ensure operational continuity and silent-failure prevention.

Data Analyst

Jun 2024 – Apr 2025

Econet Wireless Zimbabwe · Harare, Zimbabwe

- Partnered with product and strategy teams to conduct root cause analysis on consumer acquisition and growth channels, surfacing actionable insights for product decisions.
- Engineered automated **Python** validation and anomaly detection scripts to enforce strict data quality rules across operational pipelines.
- Executed advanced multi-table **SQL** queries on large relational data warehouses to isolate and report on performance metrics across business units.
- Developed dynamic reporting systems to communicate complex statistical findings and data warehouse metrics to non-technical business leadership.

Junior Data & Software Engineer — Intern

Jun – Sep 2023

Dimension Data South Africa (NTT Ltd.) · Johannesburg, South Africa

- Managed data schema integrity and logged structural errors using **SQL** during large-scale enterprise CRM migration projects.
- Validated structural integrity of legacy systems to ensure safe extension into new enterprise environments without data loss or corruption.
- Developed rigorous **SQL** extraction scripts ensuring data validation for enterprise-level reporting platforms.

Freelance Full-Stack Developer

Prior to 2023

Independent · Remote

- Delivered end-to-end web applications for confidential clients using **React, FastAPI, and PostgreSQL**.
- Managed full project lifecycle from requirements gathering through deployment and client handoff.

PROJECT EXPERIENCE

Volarix - GenAI Financial Market Stress Intelligence System

2025 – Present

Independent Research Project · Live at volarix-rouge.vercel.app · **HuggingFace: sabinalucy-volarix.hf.space**

- Architected and deployed a production-grade end-to-end ML system ingesting **870K+** deduplicated financial headlines (**Kaggle + NewsAPI**) alongside six FRED macroeconomic time series to predict market stress regimes in real time.
- Built from scratch a **PyTorch Temporal Fusion Transformer** for multi-horizon forecasting, an **XGBoost** regime classifier with Platt calibration, split conformal prediction intervals for uncertainty quantification, and a historical analogue engine.
- Deployed 9 live REST endpoints on **HuggingFace Spaces via Docker** with a **FastAPI** backend; built a **React** dashboard deployed on **Vercel** for real-time visualisation of stress signals, regime states, and forecast bands.
- Managed full 9-phase build lifecycle: data ingestion, **NLP** feature engineering, model training, calibration, **API** design, containerisation, and production deployment.

Infra Anomaly AI - Infrastructure Risk Detection via Time-Series Deep Learning

2025 – 2026

Independent Research Project · **GitHub: github.com/SabinaLucy**

- Built an unsupervised deep learning system using **Dense Autoencoder and LSTM Autoencoder** architectures to detect abnormal behaviour in infrastructure load signals - no labelled failure data required.
- Processed and modelled a dataset of over **2 million records**, managing the full pipeline from data normalisation through model evaluation, threshold selection, and anomaly flagging.
- Compared reconstruction error profiles between architectures to evaluate sensitivity to temporal dependencies and subtle signal drift in periodic load patterns.

PolyRisk - ICU Patient Deterioration Early Warning System

2025 – Present

Independent Research Project · **In Progress**

- Designing a multi-condition patient deterioration early warning system on MIMIC-IV clinical data to surface early risk signals across multiple simultaneous clinical pathways in ICU settings.
- Targeting real-time risk scoring that enables proactive clinical intervention before critical deterioration events occur.

ConstructIQ - Multimodal Agentic AI for Construction Intelligence

2026

Independent Research Project · **In Progress**

- Designing an agentic AI system combining **YOLOv8** safety violation detection, Vision Transformer blueprint analysis, **GraphRAG** document intelligence, and **LangGraph** autonomous audit pipelines for the construction industry.
- Architecture targets dual deployment on HuggingFace Spaces and AWS EC2 with a React 18 frontend on Vercel.

HAR Deep Learning - CNN vs LSTM vs CNN-LSTM Hybrid (Human Activity Recognition) 2025

Academic Group Project · Michigan Technological University

- Benchmarked three deep learning architectures (**CNN, LSTM, CNN-LSTM hybrid**) on UCI HAR smartphone sensor data across 6 activity classes with full ablation study and training dynamics analysis.
- Executed training on MTU's HPC cluster; produced comprehensive architecture comparison report.

Chronic Disease Risk Profiling - CDC Data Mining 2025

Academic Group Project · Michigan Technological University

- Applied a full data mining pipeline (**XGBoost, Random Forest, AdaBoost, Logistic Regression, KMeans, DBSCAN, PCA**) on the CDC Chronic Disease Indicators dataset to surface population-scale risk patterns.
- Performed full cross-validation, hyperparameter tuning, and cluster profiling for actionable risk segmentation insights.

Tech Shop - E-Commerce Platform Database Design 2025

Academic Project · CS3425 Database Systems · Michigan Technological University

- Designed and implemented a full relational database schema for an online store: ER modelling, normalisation to 3NF, complex multi-table SQL queries, stored procedures, and transaction management.