

THOMAS AMOGOLLA TSUMA

Machine Learning Engineer

Malaysia • +60 19 762 3483 • tommytsuma7@gmail.com • github.com/TomiTsuma • linkedin.com/in/tomitsuma • [portfolio](#)

SUMMARY

Machine learning engineer with 5+ years building production ML systems across agri-tech, fintech, and sustainability — from GAN-based spectral models that secured a USD 3M contract to credit-risk models embedded in live lending pipelines. Full-stack across data pipelines, model development, MLOps, and cloud deployment (AWS, Azure). AWS Machine Learning Specialty certified; currently completing an MSc in Data Science.

EXPERIENCE

Machine Learning Engineer — Crop Nutrition Laboratory Services (Cropnuts) *Sept 2022 – Present | Limuru, Kenya*

- Secured a USD 3M revenue contract with KALRO by designing pix2pix GAN models that generate synthetic Alpha-ATR spectra, enabling analysis of 60,000+ soil spectra.
- Built an end-to-end ANN optimization pipeline — sourcing soil spectra from PostgreSQL and wet-chemistry results from MySQL, versioning with DVC, storing on S3, and deploying to DigitalOcean — cutting model update turnaround to 4 days.
- Designed an RNN autoencoder outlier-detection system that filters erroneous spectra before inference, reducing prediction delays.
- Developed an EPO conversion workflow allowing 2mm Alpha-ATR spectra to mimic Fritsch-milled spectra, eliminating rescanning for Loyola University sample analysis.
- Built a multi-label Random Forest classifier identifying plant types and growth stages, reducing sample misclassification ahead of lab testing.
- Deployed models as Azure Functions on Azure ML and co-built a model evaluation tool with scatterplot visualizations and metrics (RMSE, R^2 , RSC, slope, kurtosis, skewness).

Mobile Developer & AI Engineer — Paylend Limited *Jan 2022 – Sept 2022 | Nairobi, Kenya*

- Cut client onboarding from ~2 days to a few hours by engineering a TensorFlow model that detects ID cards in uploads and matches faces to profile photos, flagging fraudulent accounts early.
- Built a logistic regression credit-risk model integrated into the loan approval pipeline, identifying ~60% of eventual defaulters before loan-limit increases.
- Replaced arbitrary credit decisions with data-driven loan limits via an RFM (Recency, Frequency, Monetary) customer analysis model.
- Developed an ARIMA seasonal forecasting model of loan issuance patterns that informed strategic advertising cycles.
- Launched v2 of the Paylend app — business, school-fees, and personal loans, fundraising, and utility payments — extending reach to three countries.
- Deployed inference APIs with Flask and Gunicorn on EC2 and built a Dash/Plotly BI dashboard for internal decision-making.

Machine Learning Engineer — Circularity Ltd *Nov 2020 – Jul 2022 | Ongata Rongai, Kenya*

- Built a plastic-bottle identification system for waste sorting, combining an InceptionV3 transfer-learning classifier with a TensorFlow Object Detection pipeline to offset each model's individual weaknesses.
- Delivered the MVP that led to a signed Memorandum of Understanding with Coca-Cola.
- Represented the product in a feature on KTN's The Innovator, securing national media traction.

PROJECTS

Core&Outline — End-to-End Data Automation & Intelligence Platform — Co-Founder & ML Lead

AWS Lambda/Glue, S3, Snowflake, Kafka, Flink, Spark, Flask, LLMs

- Built a low-code analytics platform automating the path from raw data to dashboards and ML insights: Lambda ingestion from Snowflake/PostgreSQL/MySQL/CSV to Parquet on S3, with Glue + Great Expectations quality checks feeding a Snowflake warehouse of business metrics (MRR, churn, CTR).
- Developed real-time session tracking with Kafka, Flink, and Spark; a reinforcement-learning recommendation engine from click/purchase data; and a text classifier routing complaints to LLM-powered agents on Hugging Face.

Rydr — Route-Based Ride Matching App — Co-Founder

React Native, Node.js, Express.js, IBM Cloud, Google Maps API

- Launched a Google Play Store app matching passengers and drivers by route similarity and departure time, backed by an Express.js API on IBM Cloud Foundry using the Google Maps API.

EDUCATION

MSc, Data Science — INTI International University, Malaysia (Expected 2026)

BSc, Computer Science (Computational Science), Second Class Upper Honours — Multimedia University of Kenya (2017–2021)

CERTIFICATIONS

AWS Certified Machine Learning – Specialty (MLS-C01) • Microsoft Azure AI Fundamentals (AI-900) • IBM Data Science Fundamentals • Huawei Certified Network Associate (HCNA)

TECHNICAL SKILLS

Languages: Python, SQL, JavaScript, Java, Dart, R, C, C++

ML & Data: TensorFlow, Keras, PyTorch, scikit-learn, NLTK, OpenCV, Spark, Kafka, Flink, DVC, Great Expectations

Web & Mobile: Flask, Django, Node.js, Express.js, React, React Native, Flutter, Android

Databases: PostgreSQL, MySQL, SQL Server, Snowflake

Cloud & DevOps: AWS (SageMaker, Lambda, Glue, EC2, S3, CodeDeploy), Azure ML, Google Cloud, DigitalOcean, IBM Cloud, Docker, Git, Swagger