

JOHNSON MUGARRA

Data Scientist | AI/ML Specialist | Research & MEAL

Kampala, Uganda · +256 777 108 405 · johnsonmugarra@yahoo.com · johnsonmugarra.com · linkedin.com/in/johnsonmugarra

PROFESSIONAL SUMMARY

Data scientist with 6+ years designing, deploying, and scaling machine-learning solutions, predictive-analytics systems, and automated data pipelines across agriculture, health, and international development. Experience spans the full ML lifecycle, from framing the problem through model development, deployment, monitoring, and retraining, alongside interactive BI dashboards and rigorous data-quality frameworks. Co-author on peer-reviewed agronomy research and comfortable translating complex analysis for technical teams and executive audiences alike, up to national policy level.

CORE TECHNICAL SKILLS

Programming & ML Python, R, SQL, Java · TensorFlow, PyTorch, Scikit-Learn, Statsmodels, Prophet, XGBoost, Keras

Methods Time-series forecasting, classification & regression, clustering & segmentation, causal inference, NLP (sentiment & network analysis), computer vision

Visualisation & BI Power BI (Microsoft Certified), Tableau (Certified Specialist), Looker Studio, Plotly, D3.js

Data & MLOps ETL pipelines, Airflow, Spark, Hadoop, Docker, Git, PostgreSQL, MongoDB, SurveyCTO, API integration, model monitoring & retraining

Cloud AWS, Azure, Google Cloud

PROFESSIONAL EXPERIENCE

Research & Data Consultant

Nov 2025 - Present

Enimiro · Remote

- Own the data platform behind a vanilla and coffee agribusiness: seven scheduled pipelines covering SurveyCTO-to-PostgreSQL ETL, NASA POWER climate data, Prophet harvest forecasts, purchase-loyalty and pollination analysis, and inspection follow-ups.
- Developed a random-forest vanilla yield-estimation model over 275 surveyed farmers (14 farm features), cutting mean absolute error from a 158 kg median baseline to 101 kg on out-of-fold validation.
- Produced interactive and static district plot maps from GPS boundaries (GeoPandas, Folium) to support agrohub planning, and automated inspection and certification reporting.
- Led standalone studies: a vanilla-farmer Living Income study (212 farmers, Anker/Fairtrade benchmarks) and the Kayunga agrohub farmer needs assessment.
- Authored concept notes on data quality, field accountability, and zonal vanilla production for management.

Research & MEAL Specialist

Jan 2025 - Present

Center for Research & Communication (CRC) · Kampala, Uganda

- Lead analytics, AI adoption, and monitoring & evaluation frameworks for data-driven programme optimisation across multi-sector projects.
- Run technical workshops and training on ML/AI, statistical modelling, Python, R, and Power BI, raising data literacy across teams.
- Supervise cross-functional research teams end to end, improving programme effectiveness by 20-30% through evidence-based decision-making.
- Establish data-governance frameworks, quality-assurance protocols, and audit-compliance systems across organisational partnerships.

Data Scientist

Jan 2024 - May 2025

MAAIF · VANEX · Makerere University · Catholic Relief Services

- Designed and deployed an end-to-end forecasting system integrating API, satellite (NASA POWER), and ground data, achieving an 18% MAE reduction and 80%+ accuracy.
- Led the full ML lifecycle - data collection, model development, validation, deployment, and monitoring - with automated retraining protocols.

- Built and automated scalable ETL pipelines with daily API synchronisation using Python and Airflow, consolidating data into a central warehouse.
- Developed interactive Tableau/Power BI dashboards giving non-technical stakeholders real-time access to insights.
- Presented findings and policy recommendations to H.E. President Yoweri Kaguta Museveni (April 2025).

Research Associate

Jun 2023 - May 2025

USAID-SITES · SSS Inc. · DLH Corporation

- Supervised digital data collection and ran SQL/Python data-quality audits, achieving 95%+ data reliability for high-stakes impact evaluations.
- Conducted facility readiness assessments for the transition to Health Information Exchange across health facilities in the Lango, Kigezi, and Busoga regions.
- Built stakeholder dashboards and automated reports that enabled 15% operational-efficiency gains.
- Engineered ETL pipelines across SurveyCTO, Power BI, and PostgreSQL with daily API automation.

Data Analyst / Research Associate

Jul 2021 - Jun 2023

Bronkar (U) Limited · Kampala, Uganda

- Designed M&E frameworks and integrated digital data-collection tools, reducing field-operations time by 25%.
- Ran advanced statistical analysis and produced evidence-based recommendations for stakeholders across East Africa.
- Led technical capacity building and staff training on analytics tools, building sustainable organisational data capabilities.

SELECTED PROJECTS

Predictive Analytics & Forecasting System

2024 - 2025

- Full-stack time-series forecasting pipeline with multi-source data integration, achieving an 18% MAE improvement; Prophet with hierarchical reconciliation, deployed to a real-time stakeholder dashboard.

Vanilla Yield Estimation Model (Enimiro)

2026

- Random-forest model predicting recorded harvest for 275 farmers; beat the median baseline (MAE 101 kg vs 158 kg) with certification status, land size, and plant count as the strongest predictors.

Farmer Data System & Plot Geomapping (Enimiro)

2026

- SurveyCTO-to-Google-Sheets ETL with photo handling, plus interactive and static district maps of farmer plots (GeoPandas, Folium) for agrohubs planning.

Vanilla Farmer Living Income Study (Enimiro)

2026

- Quantified the gap between farmer earnings and a living-income benchmark using Anker and Fairtrade reference values; a modular Python/R pipeline cleaned 485 survey columns from 212 farmers into an analysis-ready dataset and written report.

Automated Impact-Evaluation Pipeline

2023 - 2026

- Python + Airflow pipeline with automated ingestion, causal-inference modelling, and auto-generated reports, reducing evaluation cycles by 85% (from 6 weeks to under 10 days).

Social-Media Sentiment & Network Analysis (MSc capstone, MAHE)

2024

- Analysed social-media data for sentiment, and modelled the network of nodes and interactions to study how information flows between users.

Computer Vision: Gesture & Object Recognition (MSc capstone, MAHE)

2024

- Explored machine perception for robotics - interpreting hand gestures (e.g. thumbs-up) and recognising objects from a camera feed.

PUBLICATIONS

- Obua, T., Sserumaga, J.P., Tukamuhabwa, P., Namara, M., Awio, B., Mugarra, J., Tusiime, G., & Chigeza, G. (2024). Unravelling Yield and Yield-Related Traits in Soybean Using GGE Biplot and Path Analysis. *Agronomy*, 14(12), 2826. <https://doi.org/10.3390/agronomy14122826>
- Co-author (2025). Identifying Optimal Lines for Enhanced Symbiotic Performance in a Mini-Core Collection of Cowpea [*Vigna unguiculata* (L.) Walp]. *Advances in Agriculture*. <https://doi.org/10.1155/aia/2202496>

EDUCATION

M.Sc. Data Science

Sep 2022 - Oct 2024

Manipal Academy of Higher Education (MAHE) · Karnataka, India

- Coursework: Machine Learning, Predictive Analytics, Time-Series Forecasting, NLP, Big Data Analytics, Advanced Regression, Statistical Modelling, Data Visualisation.

B.Sc. Agricultural Science

Sep 2016 - May 2021

Makerere University · Kampala, Uganda

- Coursework: Biometrics & Statistics, Research Methods, Mathematics.

CERTIFICATIONS

- Google Project Management Certificate
- Microsoft Certified: Power BI Data Analyst Associate
- SAS Certified Statistical Business Analyst
- Tableau Desktop Specialist

RECOGNITION

- Selected to present analytics research findings and policy recommendations to H.E. President Yoweri Kaguta Museveni at Makerere University (April 2025).