Thabang L. Mashinini-Sekgoto

(+27) 79-5399-012 | lukhetobano@gmail.com | tmashininisekgoto.vercel.app | thabang-mashinini-0081b5b6

Summary

Applied machine learning scientist and systems engineer with 10+ years of experience building production-grade AI platforms at scale. I design and operationalise end-to-end ML systems spanning data pipelines, distributed learning, and governance that survive real-world constraints and I've led mission-critical AI transformations across telecoms, banking, education, and research, translating advanced ML into measurable enterprise impact.

Education

University of the Witwatersrand   *Johannesburg, South Africa*

PhD in Computer Science (In Progress)   *2024 – Present*

- Research focus on advancing **self-supervised learning**, **probabilistic modelling**, **computer vision** and **distributed ML systems** to reduce dependency on labelled datasets and enable scalable AI adoption in real-world, low-resource environments.
- Application domains include **environmental risk modelling**, **sensor fusion**, and **real-time decisioning systems**.

University of the Witwatersrand   *Johannesburg, South Africa*

MSc in Computer Science (Distinction)   *2018 – 2019*

- **Thesis: Learning Level Set Method by Echo State Network for Image Segmentation**
- Developed a novel recurrent neural network approach using **Echo State Networks (ESNs)** for iterative image segmentation, enabling computationally efficient training and improved boundary detection.
- Research demonstrated potential applications in **medical imaging**, **remote sensing**, and **autonomous navigation** through enhanced segmentation performance.

University of the Witwatersrand   *Johannesburg, South Africa*

BSc Honours in Computer Science   *2017*

- **Project: Wildfire Estimation Using Kernel Density Estimators**
- Applied Spatial KDE and unsupervised learning techniques to model wildfire occurrence patterns across South Africa, contributing insights for **disaster management** and **early-warning systems**.

University of the Witwatersrand   *Johannesburg, South Africa*

BSc in Computational & Applied Mathematics and Astronomy & Astrophysics *2014 – 2016*

- Completed dual majors with emphasis on **mathematical modelling**, **numerical methods**, and **astrophysical simulation**, forming the foundation for later work in machine learning and probabilistic systems.

Experience

ABSA Insurance   *Johannesburg, South Africa*

Lead Data Scientist (Acting Head of Data Science)   *Mar. 2024 – Present*

- Mandated to establish and scale **enterprise-grade data science and engineering capabilities** within ABSA Insurance, defining strategy, operating models, and delivery standards.
- Built and led the **Data Science and Engineering function** from the ground up, managing a team of **3 Data Scientists and 1 Data Engineer** to advance data-driven risk mitigation in short-term insurance.
- Partnered with senior leadership to shape **AI strategy, roadmap, and capability scaling**, ensuring alignment between technical execution and long-term business outcomes.
- Architected and operationalised **enterprise ML pipelines on Databricks**, reducing model training and scoring latency by **80–90%** and contributing to a **10% improvement in client retention**.
- Designed and deployed **geospatial flood risk models** for property insurance, improving underwriting accuracy and geographic risk visibility by **80%**.
- Stabilised and modernised the **telematics driving-behaviour scoring ecosystem** used across vehicle finance and insurance products, eliminating multi-month backlogs, improving data quality, and enabling proactive digital engagement.
- Introduced **AI governance, monitoring, and model versioning** practices to strengthen reliability, regulatory compliance, and production ML performance.
- Established enterprise-wide **ETL and ML standards** via the open-source **Ubunye-AI-Ecosystems (UAIE)** initiative, creating reusable pipelines that accelerated development velocity and supported organisational capability uplift.

Vodacom   *Johannesburg, South Africa*

Senior Data Scientist (MLOps, Streaming AI, and Optimisation)   *Nov. 2021 – Mar. 2024*

- Entrusted with designing and operationalising **real-time, ML-driven optimisation systems** to improve network resilience and operational performance across Vodacom's national infrastructure.
- Provided technical leadership across data science, engineering, and IoT teams, mentoring **7 Data Scientists** while defining delivery standards and driving platform adoption.
- Partnered with senior leadership to align AI initiatives with core operational KPIs, embedding data-driven optimisation into Vodacom's national resilience and load-shedding response strategy.
- Architected and owned Vodacom's **real-time ML and optimisation platform** for network operations, integrating streaming analytics, optimisation models, and IoT telemetry to enhance maintenance decisions and improve network reliability during load-shedding.
- Developed the ⧉ **Smart Generator Monitoring and Optimisation Engine**, enabling national-scale smart-generator allocation and optimisation contributing to an estimated **≈R1 billion in annual OPEX savings** through improved fuel efficiency, maintenance scheduling, and site prioritisation.
- Designed and implemented scalable **MLOps and streaming pipelines** using Apache Flink, PySpark, Kafka, and Kubernetes, reducing model deployment cycles by **40%** and enabling reliable real-time analytics at national scale.
- Recognised with the **Vodacom Star Award** for advancing the organisation's big data, streaming analytics, and AI capabilities.

IBM Research   *Johannesburg, South Africa*

Machine Learning Research Scientist   *Apr. 2020 – Nov. 2021*

- Recruited to translate cutting-edge machine learning research into **enterprise-ready geospatial and climate intelligence systems** in collaboration with global IBM research and engineering teams.
- Designed and operationalised **large-scale geospatial and meteorological analytics pipelines**, enabling climate risk assessment, environmental intelligence, and disaster resilience use cases for enterprise and public-sector clients.
- Developed **state-of-the-art predictive models** using **TensorFlow**, improving seasonal climate forecast accuracy by **15%** and strengthening decision-making reliability across multiple industries.
- Integrated advanced ML workflows into the ⧉ **IBM PAIRS Geospatial Platform**, expanding real-time geospatial data access and supporting strategic planning at global scale.
- Led the design and deployment of the ⧉ **COVID-19 Analytics Dashboard** for Gauteng public health authorities, delivering real-time epidemiological insights to support operational response during the pandemic.
- Provided technical and research leadership by co-leading a cohort of **3 PhD and 2 MSc research interns**, ensuring research rigor, production relevance, and collaborative delivery across multiple ML projects.
- Served as technical lead within an international, cross-functional research group, aligning research objectives with platform delivery and coordinating development across regions.
- Co-authored and published peer-reviewed research at ⧉ **NeurIPS**, advancing state-of-the-art machine learning methods and strengthening IBM's global research presence.

Business Intelligence Service (University of the Witwatersrand)   *Johannesburg, South Africa*

Data Scientist   *Jun. 2018 – Apr. 2020*

- Joined a central analytics division responsible for integrating institutional data and translating it into decision-support insights for executive leadership and government subsidy planning.
- Designed and operationalised **end-to-end machine learning pipelines** in collaboration with BI analysts and engineers, supporting student success, performance management, and strategic planning initiatives.
- Developed a **course recommendation system** using K-modes clustering, achieving a **90% silhouette score** and generating approximately **R2 million annually** in government subsidy through improved enrolment and pass-rate alignment. ⧉ **Git repository**
- Introduced structured feature engineering and model training workflows, improving model reliability and reproducibility across analytics use cases.
- Led and facilitated over **50 workshops** for staff and students, strengthening institutional capability in data science, analytics engineering, and applied machine learning.

Council for Scientific and Industrial Research (CSIR)   *Pretoria, South Africa*

Data Scientist and Software Engineer   *Nov. 2017 – Jan. 2018*

- Selected to work within a multidisciplinary team to develop **data-driven decision-support systems** for South African municipalities and public-sector stakeholders.
- Designed and implemented **predictive profiling models** using Python and Scikit-learn, achieving **75% accuracy** and identifying operational bottlenecks that improved service delivery planning.
- Developed and deployed a **Django-based decision-support web application** for municipal managers, enabling real-time data access and operational visibility.
- Recognised for innovation in applied data science and public-sector digital solutions by ⧉ **Mail & Guardian**, ⧉ **CSIR**, and ⧉ **DST**.

ZAR X   *Johannesburg, South Africa*

Systems Analyst and Support   *Jun. 2017 – Oct. 2017*

- Supported and enhanced the reliability of a high-availability financial trading platform, collaborating with IT and engineering teams to resolve critical system issues.
- Implemented performance improvements and workflow refinements that reduced platform downtime by **20%**.
- Improved reporting accuracy and operational visibility by strengthening data integrations across trading and analytics systems.

Wits Astronomy Club (WAC) *Johannesburg, South Africa*
Co-Founder and Executive *Jan. 2014 – Feb. 2017*
- Co-founded the ⧉ **Wits Astronomy Club** as a science outreach and public engagement initiative aligned with the **Square Kilometre Array (SKA)** vision.
- Designed and led public lectures, stargazing events, and workshops that connected students, researchers, and industry partners, strengthening astronomy awareness and STEM engagement.
- Secured partnerships and funding to support sustained club operations, growth, and national visibility.
- Featured in media interviews with **SKA**, **SABC News**, and **eTV**, highlighting the club's role in advancing astronomy outreach in South Africa. ⧉ **Crazy-TV**, ⧉ **SABC News**

Skills

Technical Leadership	AI/ML strategy, Data science capability building, Cross-functional leadership, Technical roadmaps, Mentorship, Agile delivery, Enterprise AI governance.
Programming	Python, C/C++, MATLAB, R, CUDA, Shell Scripting, SQL, JavaScript.
Data Engineering	Databricks, Apache Spark, Hadoop, Kafka, Flink, ETL/ELT pipelines, Feature stores, Data lakes, Serverless architectures
MLOps & Architecture	Model deployment, CI/CD (Jenkins, GitLab), Docker, Kubernetes, Infrastructure-as-Code (Terraform), MLflow, Monitoring, Versioning, Real-time ML systems.
Machine Learning & AI	PyTorch, TensorFlow, Probabilistic ML, Distributed ML (Dask, Ray, Spark), Deep Learning, NLP, Computer Vision, Explainable AI, Optimization.
Software Engineering	Microservices, REST APIs, High-performance computing, Systems design, Git-based workflows, Production-grade ML systems.
Analytics & Visualization	Power BI, React, HTML/CSS, Dash, Interactive dashboards, Data storytelling.
Domains & Expertise	Scalable ML Platforms, Real-time Decisioning Systems, AI Innovation, Telecom ML, Insurance Risk Models, Geospatial Intelligence, AI Security & Ethics.
Languages	English (C2), Sesotho (C2), Setswana (C2), Sepedi (C2), isiZulu (C2)

Publications

ML-based Probabilistic Prediction of 2m Temperature and Total Precipitation
Mohamed Akram Zaytar, Bianca Zadrozny, Campbell Watson, Daniel Salles Civitarese, Etienne Eben Vos, Thabang Michael Mathonsi, <u>Thabang Lukhetho Mashinini</u> *2022*
- ⧉ **paper**

Long-Range Seasonal Forecasting of 2m-Temperature with Machine Learning
Etienne E. Vos, **THABANG MASHININI**, Ashley Gritzma, Sibusisiwe Makhanya and Campbell Watson *2020*
- 34th Conference on Neural Information Processing Systems [NeurIPS 2020].
- ⧉ **paper poster**

Mine Worker Threshold Shift Estimation via Optimization Algorithms for Deep Recurrent Neural Network
Milka C.I.Madahana, John E.D.Ekoru, <u>Thabang L.Mashinini</u>, Otis T.C.Nyandoro *2019*
- 18th IFAC Symposium on Control, Optimization and Automation in Mining, Mineral and Metal Processing [MMM 2019]
- ⧉ **paper**

Noise Level Policy Advising System for Mine Workers
Milka C.I.Madahana, John E.D.Ekoru, <u>Thabang L.Mashinini</u>, Otis T.C.Nyandoro *2019*
- 18th IFAC Symposium on Control, Optimization and Automation in Mining, Mineral and Metal Processing [MMM 2019]
- ⧉ **paper**

Selected Projects

Smart Generator Optimisation

Vodacom, South Africa

SENIOR DATA SCIENTIST

Nov. 2021 – Mar. 2024

- **Problem:** National load-shedding created severe risks to network availability, requiring intelligent prioritisation of limited generator resources across thousands of sites.
- **Approach:** Designed a **constrained optimisation and graph-search framework** integrating real-time IoT telemetry, site metadata, and network alarms via **Kafka**, **PyFlink/Flink**, and **PySpark**. Optimisation models were implemented in **CVXPY** and operationalised through the **Analytics Engine**.
- **Impact:** Reduced generator fuel usage by **10%**, lowered operational costs by **30%**, improved network uptime by **5%**, and contributed to approximately **R1 billion in annual OPEX savings**.
- Smart Generator Monitoring

Tfilterspy — Bayesian Filtering Library

South Africa

CREATOR & PRINCIPAL DEVELOPER

2023 – Present

- **Problem:** Production ML systems in telecoms and insurance suffered from noisy, unreliable sensor and telematics data, degrading downstream decision-making.
- **Approach:** Designed and built a scalable **Bayesian filtering library** supporting Kalman, Particle, and custom filters, with a **Dask-backed** distributed compute layer for large-scale time-series processing.
- **Impact:** Enabled robust state estimation and signal smoothing for IoT, telematics, and risk analytics use cases; packaged with documentation and examples and published on **PyPI** for open-source adoption.
- tfilterspy

Ubunye AI Ecosystem (UAIE)

South Africa

FOUNDER & ARCHITECT

2024 – Present

- **Problem:** Enterprise data science teams lacked standardised ETL/ML workflows, leading to fragmented pipelines, duplicated effort, and fragile production systems.
- **Approach:** Architected a modular, **YAML-driven framework** with pluggable compute backends (**Pandas, Dask, Spark**), including declarative ETL orchestration (Ubunye Engines) and probabilistic/distributed ML components (Ubunye ML).
- **Impact:** Adopted within ABSA to establish organisation-wide ML standards, reusable pipelines, and feature stores improving delivery speed, reliability, and engineering maturity across teams.
- Ubunye-AI-Ecosystems

Kasi'Lam Digital Platforms

South Africa

CO-FOUNDER

2022 – Present

- **Problem:** Township SMEs faced barriers to digital adoption due to hosting costs, technical complexity, and lack of accessible platforms.
- **Approach:** Built scalable, zero-cost web platforms using **HTML**, **CSS**, **JavaScript**, and **React**, deployed via **GitHub Pages** to eliminate hosting overhead.
- **Impact:** Enabled local entrepreneurs to establish online presence, improve customer engagement, and build sustainable digital brands at no cost.
- View Projects

COVID-19 Visual Analytics Tool

IBM Research, South Africa

MACHINE LEARNING RESEARCH SCIENTIST

Apr. 2020 – Nov. 2021

- **Problem:** Public health authorities required real-time situational awareness during the COVID-19 pandemic to guide response strategies.
- **Approach:** Developed a real-time visual analytics platform using **Python**, **Dash**, and **Plotly**, deployed on **IBM Cloud**.
- **Impact:** Enabled rapid monitoring of infections, hospital utilisation, and hotspot trends, supporting data-driven public health decision-making.
- View Dashboard

Geospatial Analytics — IBM Environmental Intelligence Suite

IBM Research (USA, Brazil, South Africa)

MACHINE LEARNING RESEARCH SCIENTIST

Apr. 2020 – Nov. 2021

- **Problem:** Global enterprises required scalable monitoring and prediction of environmental and climate risks across industries.
- **Approach:** Built geospatial analytics workflows using **IBM PAIRS**, **Python/TensorFlow**, **Hadoop**, and **Apache Airflow** to process multi-terabyte spatiotemporal datasets.
- **Impact:** Contributed core analytical components to the **IBM Environmental Intelligence Suite**, enhancing environmental risk prediction and enterprise-scale monitoring capabilities.
- Environmental Intelligence Suite

Wits Course Recommendation System *University of the Witwatersrand*

Junior Data Scientist *Jun. 2018 – Apr. 2020*

- **Problem:** First-year students lacked personalised guidance, contributing to misaligned course selection and poor academic outcomes.
- **Approach:** Developed a **K-modes clustering-based recommendation system** using **Python**, **Scikit-learn**, and **Pandas**, with outputs integrated into **PowerBI**.
- **Impact:** Achieved a **90% silhouette score** and generated over **R2 million annually** in government subsidy through improved enrolment alignment.
- ⬈ **View Recommendation Engine**

Honors & Awards

2022	**Vodacom Star Award,** Recognised for innovation, impact, and engineering excellence in developing the national Smart Generator Optimisation system.	*South Africa*
2019	**MSc Computer Science (Distinction),** Awarded for outstanding academic performance and significant contribution to research in machine learning and big data systems.	*University of the Witwatersrand*
2014–2017	**Wits Astronomy Club — Co-Founder & Executive,** Led public STEM engagement initiatives aligned with the SKA project, promoting astronomy outreach across the university and broader community.	*South Africa*

References

PhD	**Dr. Etienne Vos — IBM Research Manager,** Etienne.vos@ibm.com	*IBM Research*
PhD	**Dr. Jacobus du Toit — Machine Learning Lead,** jacoDutoit@vodacom.co.za	*Vodacom*
PhD	**Dr. Akram Zaytar — Senior Research Scientist ,** akramzaytar@microsoft.com	*Microsoft*
BSc	**Mr. Malachy Morrow — Principal Solution Architect,** mal@malmorrow.com	*Altron & Vodacom*
MSc	**Mrs Harshila Dulabh — Team Lead,** Harshila.Dulabh@wits.ac.za	*Wits University (BIS)*