

UNARINE TIMOTHY TSHIMAUWU

Data Scientist | Machine Learning | Cloud Solutions

Gauteng, South Africa | 082 580 4035 | timothytshimauwu@gmail.com

LinkedIn: [utshimauwu](#) | Portfolio: [TimothyTshimauwu](#)

PROFESSIONAL SUMMARY

Data Scientist and Analytics professional with 5 years of engineering and analytics experience, including 3+ years delivering predictive models, risk analytics, and BI solutions across financial services and large-scale manufacturing. Strong background in Python, SQL, Azure, and Power BI with documented contributions to R83M+ in cost savings, operational improvements, and revenue retention. Experienced in end-to-end analytics pipelines, stakeholder-driven reporting, and production-grade ML architecture. Seeking Data Scientist, ML Engineer, and Business Intelligence roles in banking, finance, and operations.

CORE COMPETENCIES

- **Analytics & Statistics:** Predictive Modeling | Statistical Inference | Time-Series Analysis | Regression Analysis | Hypothesis Testing
- **Programming & BI:** Python | SQL | R | Power BI | DAX | Tableau | Excel | SAS (foundational)
- **Machine Learning:** Classification | Regression | Feature Engineering | Model Validation | XGBoost | scikit-learn | Clustering (K-Means, PCA)
- **Data & Cloud:** ETL Pipelines | SQL Server | Azure Data Factory | Databricks | AWS (EC2, S3) | Docker |
- **Soft Skills:** Project Management|Business Acumen|Stakeholder Engagement|Problem_Solving

PROFESSIONAL EXPERIENCE

Financial Data Analyst | Ruach Consulting Solutions | Pretoria, Gauteng

December 2024 - Present

- Developed SQL-based data extraction and Power Query transformation workflows processing 50,000+ monthly NGO transactions, automating previously manual reporting processes through scheduled Power BI refreshes. Reduced monthly reporting time from approximately 40 to 20 hours
- Built 12 Power BI dashboards for cash flow monitoring, expenditure tracking, and funding performance across donor streams
- Performed variance analysis on R4.5M expenditure, identifying budget leakages, delivering recommendations projected to achieve R150K annual savings (5% cost reduction).

Process Engineer (Operations Data Analyst) | ArcelorMittal South Africa | Vereeniging, Gauteng

March 2022 - November 2024

- Built Azure data engineering pipeline (Data Factory, SQL Server, Power BI) processing monthly production data. Conducted bottleneck analysis identifying 6 critical equipment constraints. Cross-functional implementation contributed to R38M in operating profit through an 8% increase in throughput and a 5% reduction in downtime.
- Conducted time-series trend analysis on 5 years of customer complaint data, identifying equipment failure patterns correlated with maintenance schedules. Findings contributed to corrective actions that retained R30M+ annual-value customer account.
- Developed a regression-based optimization model analyzing energy consumption patterns for a 1.8M tonne/year plant. Analysis identified operational adjustments achieving R15M in annual energy cost savings through reduced gas consumption

Graduate Engineer | ArcelorMittal South Africa | Vereeniging, Gauteng

March 2020 - February 2022

- Analyzed 180,000+ production and quality records using SQL and Python, supporting cross-functional initiatives that delivered an 8% throughput improvement and a 10% reduction in customer complaints
- Optimized hot rolling mill process control using 50,000+ equipment records, contributing to R15M in mechanical-loss savings through improved operational efficiency and reduced unplanned downtime

KEY ANALYTICS PROJECTS

Credit Default Risk Management System|*Personal Project*| [GitHub Repository](#)

Python | SQL | SAS | XGBoost | FastAPI | Streamlit | Power BI | AWS

- **Business Problem:** Bank needed scalable credit scoring system to reduce default rates and optimize lending decisions
- **Approach:** Developed XGBoost classification model on 1M+ loan application dataset with 45 engineered features (DTI ratio, credit utilization, payment patterns), achieving 87% precision and 82% recall through hyperparameter tuning and 5-fold cross-validation
- **Outcome:** Built FastAPI scoring service containerized with Docker and deployed to AWS EC2. Architecture demonstrates production-ready design with error handling, logging, and health checks. Projected 15% bad debt reduction based on test set performance vs baseline approval strategy

ArcelorMittal Production Analytics & Azure Pipeline|*Arcelor Mittal experience*| [GitHub Repository](#)

Python | Azure Data Factory | SQL Server | Power BI | Statistical Analysis

- **Business Problem:** 504,000 tonne/year steel plant missing production targets due to limited visibility into equipment-level bottlenecks
- **Approach:** Built an Azure data engineering pipeline using Data Factory to automate ETL, reconstructed equipment cycle times from event data, and applied Pareto analysis on downtime with Power BI dashboards
- **Outcome:** Identified six critical constraints. Methodology aligned with an ArcelorMittal initiative that delivered R38M in operating profit through an 8% throughput increase, 5% downtime reduction, and the first 40kt/month production milestone

Customer Segmentation & Marketing Personalization|*Personal project*| [GitHub Repository](#)

Python | SQL | PCA | K-Means | Decision Trees | Streamlit

- **Business Problem:** Inefficient mass marketing campaigns requiring data-driven customer segmentation for personalized product offerings
- **Approach:** Applied PCA, reducing 22 features to 8 components (85% variance retained) and K-Means clustering (silhouette score 0.68) on 250,000+ customers. Built Streamlit web application for real-time segmentation
- **Outcome:** Identified 5 actionable segments with 92% prediction accuracy, projecting 25% cross-sell improvement and 15% marketing cost reduction

FNB Customer Churn Prediction & Retention Analytics|*Personal Project*| [GitHub Repository](#)

Python | XGBoost | Power BI | Statistical Analysis | CLV Modeling

- **Business Problem:** FNB Bank 20.4% annual customer churn with no predictive visibility into high-risk customers or revenue exposure quantification
- **Approach:** Using a Kaggle-transformed South African banking dataset, built an XGBoost churn model on 10,000+ customers and 43 demographic, product, and engagement features. Developed a composite churn risk score and delivered two Power BI executive dashboards for portfolio analysis
- **Outcome:** Quantified R578.5M in at-risk CLV, achieved a 1.72x lift in churn identification versus random selection, and designed a conceptual retention strategy with a projected 1,604% ROI

EDUCATION

Post Graduate Diploma in Data Science (Cum Laude) | University of the Witwatersrand | 2025

- Specialization: Machine Learning, Programming (Python, R), Mathematical Foundations

BEng (Hons): Industrial and Systems Engineering | University of Pretoria | 2023

- Specialization: Enterprise Engineering, Optimization, Manufacturing Planning & Control

BEng: Metallurgical Engineering | University of Pretoria | 2019

- Specialization: Process Control, Calculus, Statistical Analysis

References available upon request