

DAN A. TSHISUNGU

Data Scientist | ML Engineer | MLOps | GenAI | Industrial Engineer

[LinkedIn](#) [data science portfolio](#) (In-construction) Software dev portfolio coming
dantshisungu@gmail.com +27642511891

About Me

As a Full-Stack Data Scientist, I specialize in end-to-end data solutions, from data collection and cleaning to advanced analytics and model deployment. With expertise in Python, R, SQL, machine learning, deep learning, and large language models, I design and build data pipelines, create predictive models, and deliver actionable insights that drive business decisions. Passionate about leveraging data to solve complex problems, I combine technical proficiency with a strategic mindset to create impactful data-driven solutions across industries.

Skills

- **Programming:** Python, JavaScript, R, SQL, Markdown
- **Machine learning:** Scikit-learn, tidymodels, mlr3, tidyclust,
- **Deep learning:** TensorFlow, PyTorch, Torch, Keras,
- **Big data:** Spark (Sparklyr, koalas, PySpark), Arrow, duckDB, Hadoop
- **Models:** logistic/linear regression, mixed-effects, k-means, DBSCAN, Optimisation algorithms (gradient descent, PSO, genetic algorithm, grey wolf, etc), hierarchical clustering, k-nearest neighbour, bagging, random forest, dimensionality reduction (PCA, UMAP, t-SNE, auto-encoder, LDA, etc), text mining, recommenders, SOM, NLP and LLM techniques.
- **Deployment:** Vetiver, Flask, Shiny, AWS Elastic beanstalk, Kubernetes, EKS, KServe
- **Front-End Dev:** React.js, Bootstrap, Material-UI, Shiny (R and Python)
- **Back-End Dev:** Node.js, Express, RESTful API, GraphQL, fastAPI, plumber, pin
- **Database:** SQL (PostgreSQL, MySQL), NoSQL (MongoDB). Snowflake and databricks
- **DevOps:** Docker, Kubernetes, BentoML, Github actions, Git,

Education

- **Post-graduate diploma** (Industrial Eng. - Data Science), Stellenbosch, **2023-2023**
- **Bachelor of Engineering and Technology** (Indus. Eng), TUT, **2021-2022**
- **National Diploma** (Indus. Eng), TUT, **2018-2020**

Certifications

- [The Complete Full-stack Web development bootcamp 2024](#) – Udemy
- [Introduction to LLM in Python](#) – Datacamp
- [Introduction to Shell](#) – Datacamp
- [Spoken Language Processing in Python](#) – Datacamp
- [Deep Learning Specialization](#) – Coursera (DeepLearning.Ai)
- [Natural Language Processing Specialization](#) – Coursera (DeepLearning.Ai)
- [Generative AI Engineering with LLMs Specialization](#) -Coursera (IBM)

Work Experience

December 2022 – Current

Full-stack data scientist – Freelance

- Predictive customer retention shiny app with deployed model.
- Built and deployed using docker a patient readmission model with Vetiver.
- Built a reporting Shiny dashboard that is scheduled to run every morning at 8am
- Built a customer lending model powered Databricks computational power, spark, and the tidymodels framework on Posit platform.
- Developed an ELT pipeline for a production company to assist with traceability.
- Fine-tuned speech recognition whisper for low-resource language(isiZulu)
- Built and deployed a hair classifier model using tensorflow.
- Build a clothing classifier using pretrained model using PyTorch.

June 2023 – Current

Industrial Engineer – Sinoville Fencing

- Developing and implementing the company's quality management system with QA tools, QC tools, and standardizing processes with QCP.
- Conducting the process optimization.
- Improving profitability through waste management.
- Implementing traceability of products through the build-up design of a QR system.
- Implementing a data infrastructure for production data collection and processing.
- Building dashboards and reports for upper management.
- Conducting data analysis of production, sales, and marketing data.
- Decreasing order lead time through waste elimination and value stream analysis.

Sept 2022 – May 2023

Estimator – Sinoville Fencing

- Project Estimator – Sinoville Fencing
- Estimated and delivered tender quotes to clients.
- Planned, ran, and scheduled the supply and installation of materials.
- Managers resources, contractors, and personnel for successful completion of projects

Sept 2021 – Oct 2021

Student work – Weir Minerals

- Measured assembly processes of hose building.
- Applied I.E techniques (time study, root-cause analysis, 5S, ergonomics, Pareto analysis, etc) in finding bottlenecks and improving the efficiency of processes.

Jan 2021 – Dec 2021

Industrial Engineer Research Assistant – Tshwane University of Tech.

- Assisted post-graduate students with data collection and cleaning.
- Assisted the department keep track of research progress.
- Compiled reports and presentations.
- Assisted researchers with modelling and simulations using SolidWorks, CATIA/Delmia, and Simio.

Soft Skills

- **Language: English** (fluent), **French** (native), **Swahili, Lingala, Tshiluba**
- **Communication**
- **Project management**