

# Emmanuel Sifiso Nkosi

Data Scientist

✉ [emmanuelnkossi@outlook.com](mailto:emmanuelnkossi@outlook.com) ☎ +27 69 182 3052 📍 Pretoria [in](#) LinkedIn  
[GitHub](#) [Website](#)

## Professional Summary

Data Scientist and AI Engineer with real-world experience in data science, machine learning, AI engineering, and frontend development. Proven ability to extract insights from complex data, build ML models, and deploy full-stack applications in production environments.

## Experience

<b>Wits Enterprise (University of the Witwatersrand)</b> <i>Data Scientist / AI Engineer</i>	12/2024 - Present Johannesburg
<ul style="list-style-type: none"><li>- Developing AI solutions leveraging LLMs, NLP, and Deep Learning to uncover patterns and analyze large-scale social media datasets and inform business decisions in telecommunications.</li><li>- Applying advanced unsupervised learning techniques such as UMAP for dimensionality reduction and HDBSCAN for clustering, combined with BERTopic, LLaMA2, and OpenAI models to extract, label, and interpret latent topics in unstructured datasets.</li><li>- Expanding into frontend engineering using Angular and React to productize internal tools.</li><li>- Integrating API endpoints provided by backend teams to connect frontend UI components to SQL Server via token-authenticated endpoints.</li><li>- Continuously integrating and deploying AI/ML features to a live application hosted on Azure Web Apps, in collaboration with backend and frontend teams as APIs and components evolve.</li></ul>	
<b>Telkom Corporate Centre</b> <i>Junior Data Scientist (IoT &amp; Digital Twin Focus)</i>	01/2024 - 12/2024 Pretoria(Centurion)
<ul style="list-style-type: none"><li>- Worked on IoT-integrated Digital Twin proof-of-concepts for predictive maintenance and system monitoring.</li><li>- Simulated telemetry data using AWS IoT SiteWise and 3D motor models through TwinMaker.</li><li>- Trained and deployed ML models on AWS SageMaker and EC2 to analyze and forecast equipment conditions.</li><li>- Built real-time dashboards with Grafana using data stored in S3.</li><li>- Led interns through end-to-end acoustic-modeling and machine-learning projects, driving technical direction, model development, and applied ML practices.</li></ul>	

**BCX (Business Connexion)**  
*IT Systems Development Intern (NQF 5 - Data Science)*

- Built a Speech Sentiment Analysis system to monitor customer satisfaction and agent performance in call centers.
- Performed data labeling, EDA, preprocessing (MFCCs, spectrograms), and augmentation.
- Trained deep learning models including CNNs, RNNs, and Transformers; deployed and presented the project.

01/2023 - 01/2023  
Johannesburg

**BP Printing**  
*Graphic Designer / Production Operator*

- Created print-ready designs using CorelDRAW for branding, signage, and promotional material.
- Operated equipment including vinyl cutters, sublimation printers, heat presses, and laser cutters to prepare materials for tailoring and production.

01/2022 - 01/2022  
Johannesburg

Education	
<b>University of South Africa (UNISA)</b> <i>HCert in Mathematics &amp; Statistics</i> <b>Accepted for 2026 Start</b>	2026 - 2026
<b>IT Varsity</b> <i>National Certificate in Fullstack Development NQF 5</i>	2025 - 2025
<b>Explore AI</b> <i>HCert in IT Systems Development (Specialization: Data Science)</i>	2023 - 2023
<b>Vukuzame MST School</b> <i>Matric Certificate in Science</i>	2019 - 2019

**Skills**

**Languages & Tools:** Python, SQL, Angular, Node.js, Bash

**Libraries & Frameworks:** Pandas, NumPy, Scikit-learn, TensorFlow, PyTorch, OpenAI, LangChain, React, Angular

**ML & AI:** Predictive Modeling, LLMs, RAG Pipelines, NLP, Deep Learning, CNNs, RNNs, Transformers

**Deployment & DevOps:** AWS EC2, Azure Web Apps, Render (CI/CD), Vercel, GitHub Pages, GitHub Actions

**APIs & Databases:** RESTful APIs, Swagger, SSMS (SQL Server), PostgreSQL, Vector Database(Pinecone)