

JAN MOTENE

Pretoria, Gauteng, South Africa | +27793530741
motenejo@gmail.com | [Portfolio](#) | [GitHub](#) | [LinkedIn](#)

SUMMARY

Adaptable and analytically-minded **Junior Data Scientist** and **Software Developer** with passion for uncovering insights from data. Knowledge of statistical analysis, machine learning, Natural Language Processing (NLP), and Large Language Models (LLMs). Proficient in Python, SQL, and data visualization, with hands-on experience in TensorFlow, PyTorch, and NLP packages. Capable of driving data-driven decision-making to optimize processes and outcomes.

TECHNICAL SKILLS

- **Programming Languages:** Python, R, Java, SQL
- **Database Management:** MySQL, PostgreSQL, SQL Server
- **Machine Learning Tools:** TensorFlow, Keras, PyTorch, Scikit-learn, MLFlow
- **Data Analysis and Visualization:** Pandas, NumPy, Seaborn, Matplotlib, Power BI, Tableau
- **NLP and LLMs:** Experience with NLP Packages, LLM Training, Fine-Tuning, and Deployment
- **Cloud Platforms:** Spark, Dask, Databricks, AWS, Azure, Version Control (Git)
- **Other Skills:** Statistics, Statistical Modelling, Predictive Modelling, KPI Monitoring

EXPERIENCE

Sand Technologies	Remote, RSA
Junior Data Scientist	01/ 2024 – 12/ 2024
<ul style="list-style-type: none">• Developed predictive models and recommendation systems using Machine Learning (ML).• Directed Data Engineers to optimize ETL pipelines, reducing data preprocessing time by 30%.• Conducted Exploratory Data Analysis (EDA) on a 1M+ record dataset, identifying actionable insights.• Deployed high-impact analytical solutions, implement ML workflows at scale, particularly using LLMs• Evaluated prototypes, and deployed robust models to ensure scientific analysis and business value	
Tshimologong Precinct	Johannesburg, Gauteng
Junior Software Engineer	07/ 2023 – 06/ 2024
<ul style="list-style-type: none">• Developed and maintained scalable Java and Spring Boot applications, improving system performance.• Designed and implemented RESTful APIs, reducing data retrieval time by 35%.• Optimized SQL queries and database structures, enhancing overall efficiency by 20%.• Conducted unit testing and peer code reviews, ensuring zero defects in production releases.• Teamed with cross-functional teams to enhance deployment efficiency and deliver high-quality solutions.	
Dynamic DNA	Randburg, Gauteng
System Developer Intern	05/ 2022 – 04/ 2023
<ul style="list-style-type: none">• Developed and maintained RESTful APIs for seamless data integration.• Improved unit and integration testing, achieving a 95% defect-free release rate.• Led Agile Scrum practices, improving communication and reducing project turnaround time by 15%.• Documented database schema changes, ensuring compliance across 20+ configurations.	

EDUCATION

ExploreAI Academy

Remote, RSA

Data Science Practitioner

01/ 2024 – 12/ 2024

Key Skills: Data Gathering and Manipulation, Analytical Programming, Data Visualization, Data Governance, Machine Learning, Data Mining, Model Development and Deployment

Tshwane University of Technology

Pretoria, Gauteng

Diploma in Information Technology

01/ 2018 – 12/ 2021

Key Skills: Object Oriented Programming, Database Management, Data Structure and Algorithms, System Analysis and Design, Web Application Development, REST APIs

CERTIFICATIONS

- AWS Cloud Practitioner: Amazon
- Java SE 8 Programmer 1 (1Z0-808): Oracle
- Azure Fundamentals – AZ900 and DP900: Microsoft
- National Certificate in IT – System Development Level 05: MICTSETA

KEY ATTRIBUTES

- **Project Delivery:** Collaborate with teams to deploy AI solutions from inception to business adoption.
- **Model Integration:** Develop and integrate scalable LLMs to enhance NLP capabilities.
- **Data Retrieval:** Use RAG, vector databases, and semantic searches for precise data extraction.
- **Automation:** Design workflows to improve process speed, accuracy, and reduce costs.
- **Performance Monitoring:** Develop performance metrics to evaluate and maintain AI model efficiency.
- **Regulatory Compliance:** Ensure AI models and data practices meet legal and ethical standards.
- **Experimentation:** Prototype NLP and LLM systems to evaluate performance and impact.
- **Industry Insights:** Stay updated on AI advancements to align strategies with the latest innovations.

PROJECTS

- **Height Segmentation Model (Geospatial AI)**
Built a deep learning model to predict building heights from satellite images, deployed on AWS EC2.
- **News Classifier (NLP)**
Created a text classification model using Natural Language Processing (NLP) to categorize news articles into topics (e.g., sports, politics, technology).
- **Employee Management System**
Developed a full-stack system using Java Spring Boot, React.js, and MySQL, featuring role-based access and CI/CD pipelines.

REFERENCES

Available upon request