

MNCEDISI MNCWABE

DATA SCIENTIST, DATA ANALYST, ML ENGINEER

CONTACTS



+27 71-732-9160

leemncwabe29@gmail.com

Cape Town, South Africa

ABOUT ME



Proactive Data Scientist with expertise in analytics, machine learning, and statistical modeling to solve complex business challenges. Skilled in translating technical insights for diverse stakeholders and building collaborative relationships. Strengths include adaptability, rapid learning, and a strong work ethic. Passionate about leveraging programming, statistics, and ML tools to address theoretical and practical problems.

SKILLS



Machine Learning, Data Analysis, Predictive Analytics

Large Language Models (LLMs), MultiAgent AI, Retrieval Augmented Generation (RAG)

Time Series, Data Pipelines, CI/CD, Extract Transform Load (ETL)

SQLServer, MySQL, BigQuery, PostgreSQL, SSMS, KNIME Analytics

Pandas, NumPy, Scikit Learn, Matplotlib, Seaborn, Forecasting

PySpark, Natural Language Processing (NLP)

Data Visualization, Data Manipulation, Data Cleaning

Data pre-processing, Feature Engineering, Model Building, Model Deployment

Problem Solving, Teamwork, Interpersonal Skills

Analytical Thinking, Change Management, Agile Methodologies

Event Analytics Tracking Design, Data Architecture Design

Experimentation, A/B Testing

API Development, API Integration

EDUCATION



MASTER OF SCIENCE DATA SCIENCES

University of Johannesburg
2025

WORK READINESS PROGRAM

Henley Business School
2020

BSC HONOURS STATISTICAL SCIENCES

University of the Western Cape
2019

BSC MATHEMATICAL AND STATISTICAL SCIENCES

University of the Western Cape
2018

TOOLS



TOOLS AND LANGUAGES FREQUENTLY USED

- Databricks, SQL, Python, PySpark, R
- Jupyter, VS Code, LookML
- Docker, Kubernetes, Datadog
- Google Cloud Platform, Looker, Google Data Studio, AWS
- SQLServer, MySQL, BigQuery, PostgreSQL, SSMS, KNIME Analytics
- Microsoft Azure, Azure Machine Learning
- Microsoft Power BI, Tableau, Google Suite, Microsoft Office
- SAS Base, SAS Enterprise Miner & Guide, Nexidia Analytics
- GitHub, GitLab, Jira, Confluence, Slack, Microsoft Teams
- Amplitude, Notion

CERTIFICATION



- Microsoft Azure Fundamentals, Microsoft
- Python and SQL for Data Science, DataCamp
- SAS Base, SAS

LINK



LinkedIn
<https://www.linkedin.com/in/mncedisi-lindani-mncwabe-a1b087171/>

REFERENCES



References available upon request

WORK EXPERIENCE



JUN 2024

SENIOR DATA SCIENTIST

ayoba

- Achieved 81% reduction in API response latency (from 1500ms to 280ms) and improved ML model availability by migrating recommendation models to GCP Spanner, reducing model downtime incidents from 2 per week to just 1 in 5 months (90% reduction in service interruptions).
- Achieved 30% reduction in computational costs by migrating ML model processing from BigQuery to GitHub workflows, optimizing resource utilization while maintaining the same processing capacity and improving operational efficiency through automated workflows.
- Developed and deployed multiple high-impact machine learning models including Games, Articles, Services Recommendations based on user activity and geolocation, Dormant Users Retention Prediction model, and Contact Chat Recommendations, driving increased user engagement and retention across the Ayoba app.
- Implemented automated monitoring systems through CI/CD pipelines for early detection of model degradation, ensuring continuous data quality checks and proactive identification of potential issues before impacting business operations.
- Provide technical guidance to junior team members.

SEP 2022

JUN 2024

SENIOR DATA SCIENTIST

Loop

- Enhanced driver capacity planning for on-demand clients through implementing a machine learning demand forecasting model to predict hourly orders, achieving 85% prediction accuracy rate in predicting actual orders.
- Achieved a 30% reduction in kilometres driven and average of 36 minutes reduction in time spent on the road by drivers through developing and implementing a machine learning route planning algorithm for Loop's Logistics clients.
- Implemented dashboards in Looker or Google Data Studio to drive actionable insights for On-demand and Logistics clients to enhance their daily operations and customer experience.
- Use LookML to build data models to enhance data relationships between Looker and SQL databases.
- Build data pipelines using Databricks to streamline data processing and integration, enabling analytics and actionable insights. This significantly improves data accuracy and accessibility, enhancing decision-making.

AUG 2021

AUG 2022

DATA SCIENTIST

takealot.com

- Enhanced inventory optimization by developing and implementing a time series daily sales forecasting model, achieving 80% accuracy rate in predicting actual sales.
- Optimized profitability of promotions by developing and implementing a machine learning linear regression model for strategic pricing.
- Implemented a stock optimization model to find optimal prices to clear dead stock and achieved a 60% sell-through rate.
- Implemented an RFM (Recency, Frequency, Monetary) statistical model for advanced customer segmentation.
- Built a profit margin impact tool, a tool that suggests the best or optimal price to maximize profit margin of promotional products.
- Developed a probabilistic Logistic Regression model that predicts the likelihood of a product to have a high return rate.
- Enhanced the efficiency of the planning team by saving them an estimate of 3 hours of daily manual work through automating the daily stock on hand report using Microsoft Power BI and Python.
- Used Looker or Google Data studio to automate reporting of products pricing versus competitor pricing to ensure Takealot's pricing is market fit.

JUN 2020

PRESENT

EXPLORING AI & DATA SCIENCE | PERSONAL PROJECTS

Self-Employed

- Developed a recruitment automation system that automates manual candidate search for recruitment headhunting and to rank candidates by matching job descriptions to further enhance best fit candidate matching.
- Developed an AI-powered travel planner using multi-agent systems to optimize itineraries with real-time flight, accommodation, and activity recommendations.
- Designed a multi-agent RAG system leveraging LlamaIndex to extract insights from videos, research papers, and GitHub repositories via automated workflows.
- Engineered automated pipelines to track and analyze trending AI/data science topics, helping me stay up-to-date with industry trends.
- Built a Python web app for product price comparison across retailers, aiding informed purchasing decisions.
- I author and publish blogs on Analytics Vidhya, covering topics in statistics, data science, AI applications, and technical tutorials.

FEB 2020

AUG 2021

MACHINE LEARNING GRADUATE / DATA ANALYST

MultiChoice Group

- Built a machine learning Logistic Regression model to predict over 7 million customers payback behavior and achieved 76% accuracy, outperforming manual predictions.
- Increased fraud detection rate by 7% across the business using a machine learning Robust Random Cut Forest model that correctly classified over 200 cases of fraudulent activities.
- Enhanced customer experience within the Business Operations department by improving decision-making processes, achieved through the development of actionable insights using Power BI dashboards.
- Implemented reporting dashboards to actively monitor trends and KPIs, providing vital quality analysis for contact centers.
- Enabled quicker decision making for HR & Payroll functional departments through data quality monitoring and reporting using Power BI.
- Reduced average customer call wait time and enhanced productivity of inbound call center agents through automating manual tasks using Microsoft Power Automate.
- Effectively utilized Nexidia Analytics for in-depth speech analysis, leading to significant improvements in customer engagement and resolution of customer issues.

FEB 2017

NOV 2019

TUTOR, TEAM LEAD: ADMIN STUDENT ASSISTANCE

University of the Western Cape

- Led a team of student assistants that managed to assist over 200 first year students with successful and seamless registration process.
- Provided tutorial sessions and helped third year Mathematics and Statistics students pass Statistics, R and SAS programming.