THABO NGAKANE

ngakanet@gmail.com
+27746936495
www.linkedin.com/in/thabo-n-0555906b/
https://github.com/thabongakane

>>> DATA SCIENCE | MACHINE LEARNING

MOTIVATION

I am passionate about solving business problems using Data Science & Machine Learning. I systematically & creatively use my skillset to add tangible value to the team, the business, and the end-user. I am constantly learning, and always looking to improve.

SKILLS & TOOLS

Programming: Python (Base, Pandas, Numpy, Matplotlib, Scikit-Learn, Keras), SQL, R, SAS
Machine Learning: Linear Regression, Logistic Regression, Decision Trees, Random Forest, KNN, k-means, PCA, Association Rule Learning, Causal Impact Analysis
Other: Statistics, Github, Data Visualisation, MS Office Pack, Tableau, Jupyter Notebook, AWS, Google Cloud Platform

EXPERIENCE

Insights Analyst
JULY 2019 - PRESENT
To combat increasing churn, I built a *customer churn model* using Logistic Regression in Python. Customers deemed highly likely to leave (> 75% probability) were put onto a retention programme leading to a 24% reduction in churn (vs. control group)
Using product association metrics and clustering techniques, I lead the creation of six *Neo-Genres* that represented true, customer driven categories that are used for content recommendation. This lead to an increase in "customer-on-site" time of 5%
Built a predictive model using a Random Forest in Python that estimated customer loyalty scores for customers that Company X's data agency couldn't tag (r-squared 93%). This lead to a 30% increase in customers we could analyze, and contact with promotional material.
Facilitated and lead an interactive brainstorming day for students studying Data Analytics at ABC University

Junior Analyst
JULY 2019 - JUNE 2025
Used SQL & Tableau to automate the
extraction of credit data, and create a dynamic weekly report that helped senior leadership understand and investigate trends overtime, and diagnose potential issues

---

PROJECTS

Grocery Delivery Optimization
Created & applied a Genetic Algorithm in Python to search out a near-optimal route across 10 addresses. This lead to estimated savings of up to 50% in both delivery time and fuel consumption over a route based upon transaction order alone. This approach could be utilized across many industries as a way to find more optimal solutions.

"You Are What You Eat" Customer Segmentation
Used k-means clustering on grocery transaction data to split out customers into distinct "shopper types" that could be used to better understand customers over time, and to target customers more accurately with the relevant content & promotions.

THABO NGAKANE

ngakanet@gmail.com
+27746936495
www.linkedin.com/in/thabo-n-0555906b/
https://github.com/thabongakane

>>> DATA SCIENCE | MACHINE LEARNING

EDUCATION

Nat/Dip (Info. Tech/Software Dev)
2001 - 2004 – Cape Peninsula University Of Technology
DSI (Certification)
2024 - 2025

COURSES & CERTS

DATA SCIENCE INFINITY

Actionable Key Learnings: Extracting & manipulating data using SQL. Application of statistical concepts such as hypothesis tests for measuring the effect of AB Tests. Utilizing Github for version control, and collaboration. Using Python for data analysis, manipulation &visualization. Applying data preparation steps for ML including missing values, categorical variable encoding, outliers, feature scaling, feature selection & model validation. Applying Machine Learning algorithms for regression, classification, clustering, association rule learning, and causal impact analysis for measuring the impact of an event over time. Machine Learning pipelines to streamline the ML pre-processing & modelling phase. Turning business problems into Data Science solutions.

NLP 101 (Udemy)

Actionable Learnings: Sentiment Analysis on customer reviews. This could be utilized to flag up, customer complaints to a dedicated support team, improving customer satisfaction