|  |  |
| --- | --- |
| Data Scientist/Analyst  Iman Ngwepe-Ntshibida | +27 83 588 7198 | deborah.ngwepe@gmail.com  https://www.linkedin.com/in/iman-morongwa-ngwepe-ntshibida  Pretoria, South Africa |

# Skills

* SQL (SQL Server, MySQL, SQLite)
* NoSQL (MongoDB)
* Excel, Tableau, Microsoft Power BI (Data Cleaning, Modelling, Manipulation and Visualization)
* Python (Pandas, NumPy, SciPy, Matplotlib, Seaborn, Plotly, Bokeh, Scikit Learn)
* Machine Learning Algorithms: Supervised (Regression and Classification Models), Unsupervised: (Cluster Analysis, KMeans, PCA)
* Color palette selection, teamwork, communication, research
* IDEs: Google Colaboratory, Visual Studio Code, PyCharm, Jupyter Notebook, Replit, Spider

# Projects Portfolio

New city weather prediction – Personal Project

* Sourced data of JFK Airport Weather sensor from NOOA.
* Utilized **Python** for **feature engineering** of the desired columns.
* Made predictions of tomorrow’s maximum temperature using **Prophet** model.
* Predictions were ~3 degrees off.

superstore report and dashboard– Side Hustle Project and Personal

* Used two separate superstore datasets.
* First dataset, manipulated data and created a report with **Power BI.**
* Second dataset, made an **Excel** Dashboard.

Heart Disease Prediction – Personal Project

* Utilized **Jupyter Notebook** to perform data cleaning, wrangling and exploration of the imbalanced data.
* Built three binary **classification models** and compared their metrics.
* Applied grid search **logistic regression** to produce a final model with an f1-score of 89% and recall of 91% for positive cases.

Air quality analysis – Personal Project

* Imported the csv file to SQL Server.
* Utilized **SQL** to clean and extract data.
* Formulated and answered question to find out the impact of weather conditions on air quality.

# Work Experience

Data science INTERN –The Sparks Foundation - RemoteMarch 2023 - Present

* Did a supervised learning task, to predict students scores using linear regression model.

Contributor – Omdena Mumbai Chapter – Remote March 2023 - Present

* Chosen as part of a team of 90+ contributors at Omdena, under the Mumbai Local Chapter: “Monitoring and Predicting Air Quality using Machine Learning.”

DATA ANALYTICS INtERN – Side Hustle – Remote Dec 2022

* Underwent a four-week online data analytics internship.
* Course work covered: Jobberman soft skills training, Introduction to Data Analytics, Excel, and Power BI.

# Education

CERT: Women techster data science fellowship – Tech4Dev Aug 2022 – Mar 2023

Majors: Data Analytics, Statistics, Excel, Power BI, Python, Machine Learning, Soft Skills Training

Avdip: analytical chemistry – Tshwane University of Technology Jan 2020 – Jan 2021

Majors: Analytical Chemistry Practice, Atomic and Molecular Spectroscopy, Sampling and Sample Preparation, Thermal Analysis and Kinetic Methods, Electro-Analytical Techniques, Chromatography, Synthetic Chemistry

# Certificates

Applied Data science lab – World Quant University Dec 2022– Present

* Currently in project 8 of 8.
* Course work: Building and Deploying Machine Learning Models, Dash, web API, Data Governance

Google data analytics professional certificate – Coursera Nov 2022

* Data Cleaning and Manipulation with Google Sheets, Big Query, Tableau

DATA ANALYST 1&2 certificate – Entry Level Jun – July 2022

* Excel, SQLite Online, Tableau

certificate in python – Public Policy in Africa Initiative (PPiAI) Apr 2022

* January 2022 Cohort: Learned coding using Python programming language through PyCharm IDE