Bongo Seakhoa

Data Scientist | Data Analyst
bongokosa@gmail.com | +27735907659
GitHub | **Linkedin**

Education

University of Colorado Boulder — Boulder, USA
Masters of Science Data Science — Feb 2023 - Present
Percentage: 89%

Explore AI Academy — Cape Town, South Africa
Certificate Data Science — Jan 2022 - Dec 2022
Percentage: 85%

North West University — Potchefstroom, South Africa
NSC Xcel Program — Jan 2018 - Dec 2018
Percentage: 75%

Experience

Upwork | **Freelance Data Scientist** — Remote | Apr 2022 - Present

I work with data big and small companies, doing a wide variety of tasks ranging from data science consulting to Python development. I service a very broad base of client business needs. These data needs range from data analysis (More BI-centered) to product development and big data management.

1EIGHT | **Data Analyst** — (Remote) Fortitude Valley, Australia | Jan 2023 - May 2023

- Google Cloud Platform based data integration and management
- Dashboard maintenance and creation
- Streamlit based automated reporting

Explore AI Academy | **Data Science Internship** — Cape Town, South Africa | Jun 2022 - Dec 2022

- Gaining hands on experience in applying data analysis, mining, visualization, and/or reporting skills to real-world business problems
- Build machine learning models and design new prototypes that help stakeholders to address customers preferences.
- Completing projects from start to finish using the Agile methodology.

Appen | **Internet Analyst** — (Remote) Chatswood, Australiaa | Jun 2018 - Jan 2022

- I created monthly Excel spreadsheets to manage client projects and deadlines.
- Used Microsoft Word, Google Docs, and SharePoint to store and share created documents.
- Generated Images and other materials to be used later in machine learning.
- Validated data from other freelancers for quality assurance purposes.
- Added keywords, meta descriptions, and alt image tags to increase SEO presence.
- Conducted extensive internet research to provide facts and statistics for articles and blogs.
- Conducted simple keyword research and used SEO guidelines to increase web traffic, increasing web traffic by 40% on average within six months and a 16% monthly growth rate on average.
- Created, edited, and optimized approximately 1500+ pages of evergreen content to improve the SEO ranking for specific keywords. Completed accurate and polished formatting in line with publishing standards before submission.
- Provided writing support in a fast-paced environment for a variety of public-facing materials. Composed original written material for various types of publications and submitted it for approval by the supervisor.

Skills

Programming Languages:	Python, SQL, R, DAX
Libraries/Frameworks:	TensorFlow, Scikit-learn, Keras, Pandas, Matplotlib, Streamlit, Numpy, Seaborn
Tools / Platforms:	AWS, GCP, IBM, Anaconda, Data Bricks, Tableau, Power BI, Looker
Databases:	PostgreSQL, Microsoft SQL Server, Big Query, Big Table, MySQL, SQLite, IBM Db2 Database

Projects / Open-Source

Movie Reccomender system | Link *TensorFlow, Streamlit, Python, Sklearn*

- With this project, we were aiming to construct a recommendation algorithm based on content or collaborative filtering, capable of accurately predicting how a user will rate a movie they have not yet viewed, based on their historical preferences.
- With users of the system being personalized recommendations - generating platform affinity for the streaming services which best facilitate their audience's viewing.
- We were successful in creating a Streamlit POC of the app and demonstrating both the app and analytics of user preferences.

Climate change Tweet sentiment | Link *Twitter api, TensorFlow, NLTK, XGBoost, Pandas, Streamlit*

- With this project the goal was to be able to classify the sentiment of a tweet in relation to climate change.
- We do this by using machine learning models for classification, text mining, text analysis, data analysis and data visualization.
- POC was successfully created and demonstrated to stakeholders which they loved.

Spain Electricity Shortfall + rest API | Link *Flask, Python, S*

- In this project we were tasked to model the shortfall between the energy generated by means of fossil fuels and various renewable sources - for the country of Spain.
- The daily shortfall, which will be referred to as the target variable, will be modelled as a function of various city-specific weather features such as pressure, wind speed, humidity, etc. As with all data science projects, the provided features are rarely adequate predictors of the target variable. As such, you are required to perform feature engineering to ensure that you will be able to accurately model Spain's three hourly shortfalls.
- The project was a success, and it was successfully presented to stakeholders.

Certifications

- Google Data Analytics - **Google**
- AWS Certified Cloud Practitioner - **Amazon Web Services Training and Certification**
- Data Science Foundations: Statistical Inference - **University of Colorado Boulder**
- Financial Markets - **Yale University**
- IBM Data Science Profesional Certificate - **IBM**
- Applied Data Science Lab - **World Quant University**
- Advanced Data Science with IBM - **IBM**
- Google Africa Developer Training Program - **Andela**
- Professional Data Scientist Certificate - **DataCamp**