

Ian Mbaya

+1 (616) 273 1659 - [Portfolio](#) - muthomian@gmail.com - [LinkedIn](#) - [GitHub](#)

EDUCATION

Grand Valley State University

MS. Data Science and Analytics

Maseno University

BSc. Information Technology

Grand Rapids, Michigan, USA

August 2022 - April 2024

Maseno, Kisumu, Kenya

September 2013 - December 2017

TECHNICAL SKILLS

Programming Languages: Python, R, SAS, SQL

Libraries and Tools: PyTorch, Sklearn, Pandas, Numpy, Tidyverse, Tableau, Power BI, Git

ML Architectures: K-means, KNN, RNN, CNN, Transformers

WORK EXPERIENCE

Database Administrator

Creston Industrial Sales, Michigan, USA

May 2023 - August 2023

- **Product Catalog Management:** Spearheaded the review, preparation, and successful integration of comprehensive product catalogs from various vendors into the company's databases.
- **Data Maintenance for ERP and E-commerce Platforms:** Oversaw the meticulous maintenance of essential data related to tools, costs, and prices, ensuring seamless functionality and reliability of the ERP system and e-commerce platforms.
- **Continuous Improvement and Innovation:** Continually sought ways to leveraging technology and automation to enhance database management processes and e-commerce performance.

Software Quality Assurance Engineer

Sharelookapp.com, Singapore

May 2019 - July 2022

- **Quality Assurance and Testing:** Led the development and execution of comprehensive test scripts for both manual and automated testing environments using Python, ensuring the software met rigorous quality standards.
- **Data Management:** Utilized advanced SQL techniques to extract, transform, and load (ETL) data, facilitating the seamless implementation of Sharelookapp's EdTech solutions.
- **Collaborative Software Development Support:** Used JIRA and GitHub to support development teams in analyzing design specifications and maintaining high software quality throughout the development and implementation process.
- **Documentation and Reporting:** Created detailed documentation and presentations for tracking and reporting software bugs, feature implementations, and testing outcomes.

Data Manager

Kenya Copyright Board, Nairobi, Kenya

December 2017 - November 2018

- **Data Migration Leadership:** Successfully spearheaded the migration of over 10,000 legacy system records to a new ERP system, ensuring seamless transition and minimal disruption to operations.
- **Data Analysis and Aggregation:** Utilized Microsoft Excel to conduct thorough data analysis and aggregation, identifying and resolving discrepancies to ensure the accuracy and reliability of copyright records.
- **Data Integrity and Reporting:** Played a pivotal role in maintaining the integrity and completeness of data throughout the migration process by applying best practices in data management and utilizing Excel's robust functions.

PROJECTS

- **K-Nearest Neighbor (KNN),** Built a KNN algorithm from scratch using Numpy and Pandas libraries only to predict blinded bird species for citizen science project. [GitHub](#)
- **Non-Negative Matrix Factorization,** Built an NMF algorithm for multi-modal prediction of protein markers (ADT) given gene expression (RNA). [GitHub](#)
- **Recurrent Neural Network (RNN),** Implemented an RNN algorithm using base Python and the Pytorch Library to predict handwritten digits from the MNIST dataset with a 98% accuracy. [GitHub](#)
- **Diagonal Integration,** used Pytorch Sklearn libraries to build a model for predictingBuilt an NMF algorithm for predicting given gene expression from protein markers. [GitHub](#)

EXTRACURRICULAR ACTIVITIES

- **Graduate Student Assistant** at the Van Andel Global Trade Center - October 2023 - April 2024
- **IT Student Worker** at GVSU Surplus Store November 2022 - April 2023