

Mugume Twinamatsiko Atwine.

Email: twinmugume@gmail.com
Address: Kampala/Gaborone
Phone: +26756639646/ +26771474055

PROFESSIONAL SUMMARY: -

Dynamic and results-oriented Data Scientist with over a decade of experience in computer science and specialized expertise in machine learning, deep learning, and natural language processing. Proven track record in leading transformative projects for health equity, pandemic preparedness, and decision support systems across multiple international settings. Demonstrated ability to innovate and drive project success through ethical AI practices, advanced data analytics, and effective team leadership. Skilled in deploying robust models and pipelines using Python, R, and various AI frameworks to solve critical challenges in healthcare and bioinformatics.

Experience:

Data Science Specialist

Uganda Virus Research Institute (UVRI)

Project: eLwazi Open Data Science Platform(ODSP).

Sept 2021 – To date

Project Co-Lead: Data Work Group

- **Leadership & Management:** Spearheaded a diverse team divided into specialized sub-groups focusing on Catalog Development, RedCap Integration, and FAIR Assessment, among others.
- **Strategic Oversight:** Orchestrated regular update meetings, ensuring alignment with project objectives and timely delivery. Presented progress reports to the management team, highlighting key milestones and challenges.

Project Lead: Data Portals

- **Content Curation:** Direct the creation and management of metadata collection strategies, enhancing data usability and access.
 - **Team Leadership:** Manage a cross-functional team of over five members, overseeing both technical and non-technical aspects of the project.
 - **Technical Supervision:** Guided the technical development and deployment of the data catalogue, ensuring robust infrastructure and seamless integration.
-

Core Responsibilities as a Data Scientist:

- **Data Management:** Specialized in Data Modeling, Harmonization, and Meta Data Curation, ensuring data integrity and accessibility.
- **Integration and Standardization:** Implemented RedCap Integrations and conducted FAIR Assessments to align with international data standards.
- **Collaboration:** Partnered with Uganda site investigators to tailor project deliverables to local needs and specifications.

Additional Duties:

- **Infrastructure Development:** Pioneered the establishment of site infrastructure for the eLwazi project, incorporating advanced data science tools and workflows.
- **Innovation in Data Practices:** Continuously evaluated and refined data models and harmonization tools. Contributed to the creation of new data models and relevant data standards.
- **Research & Development:** Engaged in grant writing and contributed to the enhancement of High-Performance Computing Infrastructure, facilitating high-level research capabilities.

Data Scientist | Data Analyst Consultant
Africa Center of Excellence in
Bioinformatics and Data Intensive Sciences
(ACE)
Projects: **Multiple**
Dec 2019 – to date

Data Management and Analysis

- **Data Collection, Cleaning, and Labeling:** Leveraged Python and R to efficiently prepare datasets, including merging and labeling, ensuring high-quality inputs for analysis.
- **Exploratory Data Analysis (EDA):** Performed comprehensive data explorations with advanced visualizations to uncover key insights and inform model development strategies.

Machine Learning Development

- **Machine Learning Pipelines:** Designed and built scalable machine learning pipelines, optimizing each phase from data preprocessing to model deployment.
 - **Model Analysis and Evaluation:** Conducted rigorous evaluations of machine learning pipelines to ensure optimal functionality and performance.
 - **Rapid Prototyping:** Utilized state-of-the-art AI tools to develop prototypes rapidly, accelerating innovation and testing new approaches.
-

Tools and Technologies

- **Python (Anaconda):** Primary tool for scripting, data manipulation, and machine learning applications.
- **R (R-studio):** Used for statistical analysis, modeling, and data visualization.
- **Weka (Java Workbench):** Applied for machine learning experiments, particularly in developing and testing models.

Models and Techniques

- **Ensemble Models:** Specialized in constructing ensemble models such as Random Forest and Gradient Boosting for enhanced predictive accuracy.
- **Linear Models:** Experienced with Linear Discriminant Analysis and Logistic Regression for predictive modeling and data classification.

Key Projects

- **SIDA Risk:** Developed a machine learning model to predict HIV risk among youths based on behavioral data, enhancing targeted interventions.
 - **Twitter Analytics:** Led a data mining project on Twitter to assess and quantify COVID misinformation, contributing to public health strategies.
 - **ChatBot Development:**
 - **RASA Chatbot:** Created a chatbot to answer COVID-19 related FAQs, improving information accessibility online.
 - **HEAL Project:** Implemented an LLM-based chatbot to convert IDSR documentation into chat format, enhancing user engagement and information dissemination.
 - **CM Project:** Pioneered a predictive model for Cryptococcal Meningitis in the HIV-positive population, significantly aiding in early diagnosis and treatment planning.
-

Assisting Lecturer Big Data Analytics
Africa Center of Excellence in
Bioinformatics and Data Intensive
Sciences (ACE)
Project: **BRecA**
Position Type: Apprenticeship
Jan 2020 – to date

Academic Teaching and Supervision in Bioinformatics

Curriculum Development and Instruction

- **Lecture Planning:** Designed comprehensive lecture plans for Ph.D. and Master's programs in Bioinformatics, focusing on modern computational methods and theoretical foundations.
- **Course Instruction:**
 - Delivered engaging and informative sessions on **Linear Algorithms in Machine Learning**, exploring foundational techniques and their applications.
 - Taught advanced concepts in **Non-Linear Algorithms**, emphasizing their importance in complex problem solving.
 - Conducted detailed lectures on **Tree-Based Algorithms**, demonstrating their effectiveness in data classification and regression tasks.

Student Assessment and Engagement

- **Assessment Planning:** Developed and administered assessments to evaluate student comprehension and application of bioinformatics algorithms.
- **Creative Teaching Strategies:** Innovated teaching methods by integrating practical competitions such as Kaggle challenges, enhancing learning through real-world data science tasks.

Mentorship and Supervision

- **Student Supervision:** Actively supervised students during their class projects, providing guidance and support to foster research and technical skills.
- **Project Oversight:** Ensured projects were aligned with course objectives and provided critical feedback to promote academic and professional growth.

Data Scientist | Natural Language Consultant
Infectious Diseases Institute (IDI) - ACADEMY
Project: **Decision Support System.**
Nov 2021 – Oct 2023

Data Science and Analytics

- **Data Staging and Curation:** Orchestrated comprehensive data preparation phases, ensuring data integrity and readiness for analysis.
 - **Exploratory Data Analysis (EDA):** Conducted in-depth exploratory analysis with a focus on visual storytelling to uncover patterns and insights.
 - **Machine Learning Ethics Evaluation:** Implemented ethical frameworks and guidelines to ensure the
-

responsible use of AI and machine learning technologies.

Machine Learning Development

- **Machine Learning Pipelines:** Engineered and optimized end-to-end machine learning pipelines, from data ingestion to model deployment.
- **Model Tuning and Evaluation:** Applied rigorous testing and tuning methods to enhance model performance and accuracy.
- **Model Integration and Deployment:** Streamlined the deployment of machine learning models into production, ensuring scalability and stability.

Support and Enhancement

- **Prototyping and Experimentation:** Led rapid prototyping sessions to explore innovative solutions and enhance existing systems.
- **Technical Advisory:** Provided strategic advice on selecting the most effective technical methods to achieve project goals.
- **Front-end Support:** Assisted in front-end development, ensuring seamless integration of machine learning models with user interfaces.

Tools and Technologies

- **Programming Languages and Environments:**
 - Python (Anaconda): Utilized for scripting, data manipulation, and machine learning tasks.
 - R (R-studio): Employed for statistical analysis and visualization.
 - Weka (Java Workbench): Used for machine learning experiments and model development.

Models and Techniques

- **Ensemble Models:** Expertise in developing and deploying robust ensemble methods like Random Forest and Gradient Boosting to improve prediction accuracy.
 - **Linear Models:** Skilled in implementing linear discriminant analysis and logistic regression for data classification and predictive analytics.
-

Data Scientist | Data Science Team Lead
Africa Center of Excellence in
Bioinformatics and Data Intensive
Sciences (ACE)
Collaboration: Infectious Diseases
Institute, Makerere University Uganda
Objective: COVID-19 Chabot
July 2021 – Oct 2023

Project Overview: Chatbot Development Using RASA

Research and Planning

- **Technology Selection:** Conducted thorough research to identify the most suitable technologies for building the chatbot, ultimately selecting the Rasa framework for its flexibility and robustness.

Data Preparation and System Configuration

- **Data Staging and Curation:** Formatted data to align with the chosen methodologies, ensuring optimal compatibility and performance.
- **Infrastructure Setup:** Configured servers for SSH access, enhancing security and accessibility for user interactions.

Machine Learning Development

- **Building Machine Learning Pipelines:** Strategically selected components of the ML pipeline to optimize the chatbot's learning and interaction capabilities.
- **Model Analysis and Evaluation:** Regularly analyzed and evaluated the machine learning pipelines to ensure they functioned correctly and efficiently.
- **Prototyping and Experimentation:** Implemented a cycle of prototyping and rigorous testing to refine the chatbot's functionality and user experience.

Optimization and Deployment

- **Model Tuning and Evaluation:** Fine-tuned models to improve accuracy and responsiveness of the chatbot.
- **Model Integration and Deployment:** Seamlessly integrated and deployed the chatbot into the existing digital infrastructure, ensuring smooth operation and scalability.
- **Containerization and Scaling:** Employed K3S for containerization and scaling the application, facilitating manageable growth and resource efficiency.

Front-End and User Interaction

- **Front-End Building Support:** Collaborated closely with front-end developers to ensure the chatbot's interface is user-friendly and effective.
 - **User Experience Enhancements:** Focused on improving the interactive experience, making the
-

chatbot more intuitive and engaging for users.

Tools and Technologies Used

- **Programming and Frameworks:**
 - Python (Anaconda) for backend development.
 - Rasa and Botpress frameworks for chatbot development.
 - JavaScript for front-end support.
 - K3S for container management and orchestration.

Models and Techniques

- **Advanced NLP Models:**
 - Utilized Transformer models such as the DIET Entity Classifier for understanding and classifying user input.
 - Integrated SpaCy language models to enhance natural language understanding capabilities.
 - Employed tokenizers and featurizers to process text and extract features effectively.
 - Trained and refined entity recognizers to accurately identify and respond to user-specific data points.

Dean School of AI (Volunteer)
School of Artificial Intelligence
(NGO)
Online: www.theschool.ai
Premises: **McKinnel Knowledge
Center IDI**
June 2018 – Dec-2019

Educational Leadership and Instruction

Teaching and Curriculum Development

- **Instructional Expertise:** Provided high-level education in data science, employing the latest methodologies and tools such as Python and R, to ensure students are equipped with current and applicable skills.
- **Curriculum Preparation:** Dedicated to studying and preparing comprehensive class materials that align with cutting-edge industry practices, enhancing the learning experience.

Administrative and Organizational Roles

- **Class Coordination:** Managed the logistical aspects of class schedules and student engagements, ensuring smooth course delivery and effective time management.
 - **Event Organization:** Arranged and coordinated meetings and venues for student activities, facilitating enhanced learning opportunities and community
-

engagement.

Strategic Alignment and Resource Management

- **Vision Alignment:** Kept abreast of the NGO's evolving goals and the broader vision of the educational institution, ensuring that course offerings and content remain relevant and forward-thinking.
- **Knowledge Management:** Developed and maintained a robust knowledge base for students, providing them with essential resources and learning materials.
- **Interdepartmental Coordination:** Acted as a liaison between the Infectious Diseases Institute (IDI) and the School of Artificial Intelligence, aligning educational strategies with organizational objectives.

Database Administrator
Infectious Diseases Institute (IDI)
ECSACOP - NGO
Kampala Makerere
Mar 2018 – Nov 2018

Data Preparation and Management

- **Data Cleaning:** Utilized spreadsheets, R, and Python to clean and preprocess data, ensuring accuracy and usability for analysis and reporting.
- **Data Integration:** Skilled in uploading and organizing data from diverse sources to consolidate information into a unified database system.

Database Development and Customization

- **Database Setup:** Configured and managed the SUMAC database system, tailoring it to meet specific organizational needs.
- **Customization:** Developed custom modules and functionalities within the database to cater to the unique requirements of the NGO, enhancing operational efficiency and data accessibility.

Technical Support and Maintenance

- **Online Integration Support:** Provided technical assistance for integrating the database with online platforms, facilitating seamless data access and interactions.
 - **Database Maintenance:** Conducted regular database backups and performed maintenance tasks to ensure data integrity and system reliability.
 - **Webmail Management:** Managed webmail systems, overseeing the setup, maintenance, and troubleshooting to ensure smooth email communications within the organization.
-

Data Analysis and IT Consultation

- **Data Analysis and Reporting:** Performed detailed data analysis, extracting key insights and compiling comprehensive reports to support decision-making processes.
- **IT Consultation:** Offered expert advice on internet technologies, recommending solutions based on the specific needs of the organization to optimize their digital infrastructure.

Associate Technical Engineer
Global Enterprises International
(Malaysia) Sdn Bhd
Kuala Lumpur, Malaysia
Nov 2013 – Dec 2014

High-Volume Technical Support

- **Issue Resolution:** Provided rapid technical support to the engineering team, efficiently resolving up to 40 issues per day, ensuring minimal downtime and maintaining productivity across executive projects.

Problem Analysis and Troubleshooting

- **Technical Investigations:** Conducted thorough investigations of complex technical problems, identifying root causes related to both software and hardware issues using advanced diagnostic tools over the internet.
- **Corrective Actions:** Initiated and implemented corrective measures to address and resolve issues, utilizing online tools for tasks such as IP management and troubleshooting blacklisted sites.

Product and Software Management

- **Product Assessment:** Evaluated product issues to determine disposition and identify related technical challenges, dealing with both internal systems and third-party software across diverse platforms including Mac and Linux.
- **Collaboration and Escalation:** Worked closely with second-level technical and network teams to ensure swift resolutions, maintaining an 80% Service Level Agreement compliance and enhancing overall service delivery.

Cross-functional Team Collaboration

- **Team Liaison:** Acted as a key point of contact between first-level support and higher-level technical teams, facilitating effective communication and quick problem-solving strategies to uphold service standards.
-

Computer Technician

Ultimate Computer Solutions Ltd
Mbarara, Uganda
Feb 2009 – July 2011

Software Installation and System Maintenance

- **Software Installation:** Proficiently installed and configured various software systems including Windows, Linux, and Microsoft Office, ensuring optimal performance and user compliance.
- **Hardware Maintenance:** Conducted extensive hardware repairs and maintenance tasks, including RAM installations, motherboard replacements, and processor upgrades on over 100 computers monthly, enhancing system reliability and performance.

Sales and Account Management

- **Financial Oversight:** Managed sales accounts, effectively handling financial transactions and maintaining Excel databases for revenues exceeding 5 million shillings per month.
- **Data Management:** Ensured accurate and timely recording and reporting of financial data, supporting financial planning and decision-making processes.

Leadership and Part-time Management

- **Operational Management:** Oversaw part-time management responsibilities, coordinating daily operations and assisting with the strategic planning and execution of business activities.
- **Team Coordination:** Directed and supported team efforts, optimizing productivity and ensuring alignment with organizational goals.

•

Education:

**Master of Computer Science, Area:
Computer Networks**

Limkokwing University of Creative
Technology Malaysia
Cyberjaya, Malaysia
2013

- Honors, First Class

**Cert: Big Data Theory and Computational
Thinking** University of South Australia
(Adelaide-X)
Adelaide, Australia

- 68%
-

2017

Bachelor of Science in Computer Science

• Second class

Mbarara University of Science and
Technology
Mbarara Uganda
2011

Languages

French English

Intermediate Advanced

Honors & Awards

Silver Award, MARS Malaysian Research Society, MCAR- Mobile cloud, Augmented

Reality Silver Award, MARS ITEX, MCAR- Mobile cloud, Augmented Reality

Publications

1. Analysis of Cloud Computing Models Based on Common Factors, 2012, Atwine Mugume Twinamatsiko, Ali Naser, Jugal Joshi, Behrang Parhizkar.

1.Dynamic Workload Performance Optimization Model For Multiple -Tenancy In Cloud Systems Atwine Mugume Twinamatsiko, Jugal Harshvadan Joshi, Ali Naser Abdulhussein Abdulhussein, Arash Habibi Lashkari, Mohammad Sadeghi, IJCSI journal. Volume 10 Issue 6 Nov 2013

1.Jugal Harshvadan Joshi, Atwine Mugume Twinamatsiko, Ali Naser Abdulhussein Abdulhussein, Arash Habibi Lashkari, Mohammad Sadeghi, An enhanced Quality model for IaaS provider with non-quantifiable Key Performance Indicators, 3rd International Conference on Information Computer Application (ICICA 2014), Barcelona, Spain

Ali Naser Abdulhussein Abdulhussein, Jugal Harshvadan Joshi, Atwine Mugume Twinamatsiko, Arash Habibi Lashkari, Mohammad Sadeghi, An Efficient Load Balancing Algorithm for virtualized Cloud Data Centers, The 2014 International Conference on Circuits, Systems, Signal Processing, Communications and Computers (CSSCC14), Italy, 2014

Certificates.

Data Scientist with Python

Issuer: DataCamp

Site: <https://www.datacamp.com/>

Certificates obtained:

- Introduction to Python
 - Intermediate Python
 - Data Manipulation with Pandas
 - Exploratory Data Analysis
 - Unsupervised Learning with Python
 - Merging Data with Pandas
 - Statistical Thinking in Python part 1
 - Statistical Thinking with Python 2
 - Python Data Science Toolbox Part 1
 - Python Data Science Toolbox Part 2
 - Cleaning Data with Python
 - Introduction to Data Visualization with Matplotlib
 - Supervised Learning with Scikit Learn
 - Machine Learning with Tree-based models in Python
 - Cluster Analysis in Python
 - Introduction to Data Visualization with Seaborn
 - Introduction to importing data with python
 - Intermediate importing data in Python
 - Intermediate data visualization with Seaborn
 - Working with Dates and times in Python
 - Analyzing Police Activity with Pandas.
-

Data Analyst with SQL

Issuer: Data Camp

Site: <https://www.datacamp.com/>

- Introduction to SQL Server
- Introduction to Relational Databases in SQL
- Intermediate SQL Server
- Time Series Analysis in SQL Server
- Functions for Manipulating Data in SQL Server
- Database Design
- Hierarchical and Recursive Queries in SQL Server
- Transaction and Error Handling in SQL Server
- Writing Functions and Stored Procedures in SQL Server
- Building and Optimizing Triggers in SQL Server
- Improving Query Performance in SQL Server

Machine Learning Scientist:

Issuer: DataCamp

Site: <https://www.datacamp.com/>

- Linear Classifiers in Python
- Extreme Gradient Boosting in XGBoost
- Dimensionality Reduction in Python
- Pre-processing for Machine Learning in Python
- Machine Learning for Time Series Data in Python
- Feature Engineering for ML in Python
- Model Validation in Python
- Natural Language Processing in Python.
- Feature Engineering for NLP in Python.

Other

Issuer: DataCamp

Site: <https://www.datacamp.com/>

- Introduction to Tensorflow in Python
 - Introduction to Git
 - Experimental Design in Python
 - Introduction to Statistics in Spreadsheets.
 - Machine Learning for Business.
-