

# Hewitt Tusiime

Dallas, TX | +1 (641) 233-2632 | [tusiimehewitt@gmail.com](mailto:tusiimehewitt@gmail.com) | [linkedin.com/in/hewitt-tusiime](https://www.linkedin.com/in/hewitt-tusiime)

## Summary

---

Data Scientist & ML Engineer with 5+ years of experience applying machine learning and data automation to drive operational efficiency. Skilled in Python-based workflows using Pandas, PyTorch, and Scikit-learn, with hands-on experience building and deploying predictive models and automated data pipelines. Published researcher in NLP and multilingual AI, adept at translating complex datasets into scalable solutions.

## Technical Skills

---

- **Languages:** Python, Java, SQL, R, JavaScript, TypeScript
- **ML & Data Science:** Scikit-learn, PyTorch, TensorFlow, XGBoost, LightGBM, Statsmodels, Time Series Analysis, Data Analysis, Machine Learning, Neural Networks
- **NLP & LLMs:** Transformers, BERT, GPT, LLMs, RAG, Embeddings, spaCy, NLTK, Hugging Face
- **Computer Vision:** CNNs, OpenCV, Image Classification, Object Detection, Feature Extraction
- **Data Engineering:** Pandas, NumPy, PySpark, Apache Airflow, dbt, ETL/ELT Pipelines
- **MLOps & Deployment:** MLflow, TensorFlow Serving, TorchServe, Model Versioning, A/B Testing
- **Backend Development:** Python, FastAPI, Flask, Django, Java, Spring Boot, REST APIs, Microservices
- **Databases:** PostgreSQL, MySQL, MongoDB, Redis, FAISS, Pinecone, ChromaDB
- **Big Data & Streaming:** Apache Kafka, Spark, Distributed Computing, Event-Driven Architecture
- **Cloud & Infrastructure:** AWS, SageMaker, EC2, Lambda, S3, RDS, Docker, Kubernetes, Terraform
- **Visualization & BI:** Matplotlib, Seaborn, Plotly, Tableau, Power BI, Grafana
- **Tools & Practices:** Git, Jupyter, MLflow, DVC, CI/CD, Agile/Scrum, Experiment Tracking
- **Soft Skills:** Analytical Thinking, Independence

## Professional Experience

---

### Makerere University, Centre for AI & Data Science

Aug 2019 - Jun 2024

#### Data Scientist & ML Engineer

- Designed and deployed 10+ machine learning models including classification, regression, NLP, and computer vision systems serving 50K+ daily predictions with 99.9% uptime.
- Built end-to-end ML pipelines from data ingestion and feature engineering to model training, validation, deployment, and monitoring using Python, PyTorch, and Scikit-learn.
- Developed production LLM-powered applications using GPT, BERT, and custom transformer models for text classification, sentiment analysis, and multilingual NLP across 5+ African languages.
- Implemented Retrieval-Augmented Generation (RAG) systems with vector embeddings and semantic search, leveraging neural networks and improving information retrieval accuracy by 45% over baseline models.
- Created comprehensive data preprocessing workflows handling 2M+ records daily including cleaning, normalization, feature extraction, and data validation using Pandas and PySpark.
- Implemented MLOps practices with MLflow for experiment tracking, model versioning, and automated retraining pipelines independently, reducing model deployment time by 60%.
- Conducted statistical analysis and A/B testing to evaluate model performance, identify key features, and deliver insights through analytical thinking that led to targeted model improvements and measurable business impact.
- Developed backend microservices with Java Spring Boot and Python to serve ML models, manage data pipelines, and integrate with enterprise systems, which improved system reliability and shortened inference latency.
- Collaborated with researchers, engineers, and domain experts to define ML requirements, validate models, and deploy solutions, resulting in faster project delivery and alignment with business goals.

### Jumia Group (E-commerce Platform)

Jul 2018 - Jul 2019

#### Data Analyst

- Built machine learning predictive models for customer churn, lifetime value estimation, and demand forecasting using logistic regression, random forests, and gradient boosting, reducing churn by 18%.
- Performed exploratory data analysis and statistical hypothesis testing on customer behavior, conversion funnels, and transaction patterns informing product and marketing decisions.
- Created interactive dashboards and visualizations using Tableau and Plotly enabling stakeholders to track KPIs and business metrics in real-time.
- Implemented recommendation algorithms and customer segmentation models using collaborative filtering and clustering techniques improving campaign targeting effectiveness by 25%.
- Collaborated with engineering teams to implement A/B testing frameworks and experiment tracking for data-driven product development.

## Technical Projects

---

### AI Ticket Classification & RAG System

- Built a microservices-based ticketing system in Python using Spring Boot and FastAPI, leveraging Hugging Face transformer models (PyTorch) for ticket classification with 89% accuracy and a RAG pipeline for automated responses; deployed on AWS with Kafka streaming, auto-scaling, and monitored model performance.

### Depression Detection from Speech

- Developed a convolutional neural network audio classifier in Python using MFCCs and spectrograms for feature extraction (87% F1), implemented data preprocessing and augmentation pipelines with NumPy and Pandas, and deployed the model as a FastAPI service on AWS for real-time inference.

### **Multilingual Sentiment Analysis System**

- Fine-tuned multilingual BERT on 2M+ social posts across multiple African languages. Built data collection pipelines and deployed a Flask API handling batch inference and daily traffic with monitoring dashboards.

### **Document Intelligence (RAG) System**

- Built an end-to-end RAG workflow with document parsing, embeddings, and FAISS vector search over 100K+ chunks. Integrated LLMs with prompt engineering to deliver accurate, grounded responses from internal documents.

### **Education**

---

#### **Maharishi International University**

*Master of Science, Computer Science*

**Apr 2027**

*Fairfield, IA, USA*

#### **Makerere University**

*Bachelor of Science, Software Engineering*

**2017 - 2022**

*Kampala, Uganda*

### **Publications & Research**

---

- AI Letters (2024). *Building Text and Speech Benchmark Datasets for East African Languages*
- African NLP @ EACL (2021). *Misinformation Detection in Luganda-English Code-Mixed Text*