IAN KORIR
DATA SCIENTIST

Phone: +254 799179690 | Email: iantraven13@gmail.com | Address: Nairobi, Kenya
Portfolio: https://sirian71.github.io/ | LinkedIn: iankorir

PROFESSIONAL SUMMARY

Passionate data science and AI enthusiast, proficient in Python, TensorFlow, NLP, and Linux systems. Developed a Swahili speech emotion recognition (SER) system, achieving 87.8% accuracy using CatBoost. Operationalized the system through Streamlit and FastAPI for call center analytics. Engineered the Regan AI agent using Gemma3 LLM and PostgreSQL. This system automates API calls for storing and retrieving user memories in a database with the Telegram API and n8n workflows, seamlessly integrating with Supabase for cloud services. Certified as a Linux System Administrator (LPIC 1/2), with expertise in Apache, MySQL, and network security, experienced in end-to-end ML pipelines, data preprocessing, feature extraction, and model evaluation. Driven by a passion for ethical AI, I aim to apply NLP and predictive analytics skills to develop socially impactful real-world solutions in collaborative environments.

SKILLS

- **Programming:** Python, SQL
- **Data Visualization:** Tableau, seaborn, Google Sheets
- **ML Tools:** TensorFlow, Regression analysis, Jupyter Notebook, Pandas
- **Deployment:** Streamlit, FastAPI, n8n, supabase
- **Systems:** Linux, ngrok, PostgreSQL
- **Specialized skills:** AI Agents Integration.
- **Soft Skills:** Leadership, Teamwork, Problem-solving, Critical thinking, Time management

WORK HISTORY

Junior Linux Systems Administrator   **MAY 2022 - JULY 2023**
Linux Learning Centre | Hurlingham, Nairobi
Reporting to Senior Linux Administrator,

- Successfully executed a range of crucial tasks aligned with LPIC 1 and LPIC 2 certifications, contributing to the seamless functioning of the organization's IT infrastructure.
- Visited Local high schools like Mangu High School to spread knowledge on the Linux operating system and open source software.
- Mentored Lenana School computer students and briefly educated them on the introduction to Linux and its commercial use cases.
- Created marketing posts on social media to advertise courses that were offered.
- Installed and tested different Linux distributions on various computers and analysed their performance.
- Successfully executed a range of crucial tasks aligned with LPIC 1 and LPIC 2 certifications, contributing to the seamless functioning of the organization's IT infrastructure.
- Demonstrated proficiency in file system management, employing tools to handle file systems and partitions

- Installed, configured, and secured critical services including Apache for web hosting, MySQL/PostgreSQL for databases, and Postfix for mail services.
- Implemented robust security measures, including firewall configurations to secure file transfers via SSH.
- Created course content using LibreOffice.

Industrial Trainee   **APRIL 2021 - JUNE 2021**
Masinde Muliro University

Trained in the following fields
- Database management and administration.
- Technical support, such as cabling.
- System management and administration.
- Documentation.
- Network security and administration.
- OS installation

EDUCATION

Masinde Muliro University   **2019 - 2023**
Bachelor of Science in Computer Science

- Attained 2.2 (second class lower)
- Member of Machine Learning Club.

CERTIFICATIONS
- **Data Science**   **2024**
  - Moringa School
- **Linux Network Engineer (LPIC 2)**   **2023**
  - Linux Learning Centre
- **Linux System Administartor (LPIC 1)**   **2020**
  - Linux Learning Centre

PROJECTS

Swahili Speech Emotion Recognition System || link
Technologies: Python, Jupyter Notebook, Streamlit, FastAPI, Random Forest, accuracy metrics.

Developed a machine learning-based Speech Emotion Recognition (SER) system tailored for Kenyan Swahili, aimed at enhancing customer service interactions in call centers.

Key Responsibilities:
- Led a team of researchers in the collection and analysis of acoustic features from Kenyan Swahili speech samples.
- Implemented data preprocessing, feature extraction, and model training pipelines using Python and Jupyter Notebook.
- Evaluated multiple classification models, including K-Nearest Neighbors (KNN), Random Forest, SVM, XGBoost, CatBoost, and Multi-layer Perceptron (MLP).
- Achieved an accuracy of up to 87.8% with the CatBoost model, optimizing emotion detection accuracy.
- Deployed the SER system using Streamlit for the web application interface and FastAPI for backend API integration.
- Conducted thorough model evaluation using accuracy metrics, confusion matrix, and feature importance analysis.

Regan AI Agent (In Progress) || link

Technologies: Python, Gemma3 LLM, PostgreSQL, Telegram API, n8n, Supabase

- Automates API calls for storing and retrieving user memories in a database.
- Developing an AI-powered memory bot using Gemma3 LLM to analyze and store user information and conversational patterns in PostgreSQL via Supabase.
- Fine-tuning the model to contextualize user interactions.
- Automating workflows via the n8n low-code platform to integrate the Telegram API for real-time messaging and data updates.
- Targeting full deployment by Q4 2024, with plans to expand into personalized recommendation features.

REFERENCES

References are available upon request.