IAN KIBET KORIR

DATA SCIENTIST | DATA ANALYST

CONTACT

Phone:
0799179690

Email Address:
iantraven13@gmail.com

Website
https://sirian71.github.io

SOFT SKILLS

- Teamwork
- Time Management
- Leadership
- Effective Communication
- Critical Thinking

TECH SKILLS

- Python programming
- Data Structures
- Algorithms
- SQL, MySQL
- Data Cleaning
- Data analysis
- Machine Learning
- Linux Navigation
- System Administration

CERTIFICATIONS

Data Science

Data Science certification offered by Moringa School.

Linux Network Enginner

LPIC 2 - Certification course offered by Linux proffesional Institute.

PROFILE

I am a dynamic Data Scientist with a passion for leveraging data to drive impactful insights and solutions. Skilled in machine learning, data analysis, and system administration, with a proven track record of managing databases and deploying models effectively. Enthusiastic about contributing to innovative projects that align with impactful missions and excited to bring analytical expertise and a collaborative mindset to foster data-driven success.

EDUCATION

Data Science — 02/ 2024 - 11/2024
Moringa School

Bachelor of Computer Science — 2019 - 2023
Masinde Muliro University

Kapsabet High School — 2015 - 2018
Attained mean grade of B

EXPERIENCE

Junior Linux Systems Administrator — 05/ 2023 - 07/2023

Linux Learning Centre — Nairobi, Kenya

I 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.

Industrial Trainee — 04/ 2021 - 06/2021

Masinde Muliro University — Kakamega, Kenya

I was trained on the following fields
- Database management and administration.
- Technical support such as cabling.
- System management and administration.
- Documentation.
- Network security and administration.
- OS installation

PROJECTS

Swahili Speech Emotion Recognition System

Implemented a machine learning model for detecting emotions in Swahili language audio samples, achieving a classification accuracy of 82%. Leveraged natural language processing and deep learning to classify emotions from voice recordings, filling a critical gap for Swahili-speaking regions. The project involved data preprocessing, feature extraction, and model optimization to reliably identify different emotional states.

Technologies Used: Python, TensorFlow/Keras, Librosa (for audio feature extraction), Natural Language Processing (NLP)

Link: https://github.com/SirIan71/Speech-Emotion-Recognition-swahili

MovieLens Recommendation System

Built a recommendation system using the MovieLens dataset to provide personalized movie suggestions based on user preferences and viewing history. Implemented collaborative filtering and matrix factorization to improve recommendation relevance.

Technologies Used: Python, Pandas, Collaborative Filtering, Matrix Factorization

Link : https://github.com/SirIan71/MovieLens-Recommendations-System

House Prices Prediction Model

Created a robust model to predict housing prices based on key features such as location, property size, and amenities. Focused on data preprocessing, feature selection, and model optimization to improve prediction accuracy, aiding real estate valuation insights.

Technologies Used: Python, Scikit-Learn, Pandas, Feature Engineering, Regression Analysis

Link : https://github.com/SirIan71/House_Prices_Modeling

Chicago Crime Data Analysis and Prediction

Developed a machine learning model for analyzing and predicting crime trends in Chicago, achieving an accuracy rate of 87% in predicting high-crime areas. Employed data cleaning, feature engineering, and clustering techniques to identify patterns and forecast potential crime hotspots, contributing insights for proactive law enforcement.

Technologies Used: Python, Scikit-Learn, seaborn , Clustering

Link : https://github.com/SirIan71/Chicago_crime_data_Ml