**PROJECT PROPOSAL**

**LEVERAGING MACHINE LEARNING TO COMBAT BUSHMEAT POACHING AND ILLEGAL LOGGING IN ARABUKO-SOKOKE FOREST RESERVE, KENYA**

Primary Researchers

1. Philip Mweri
2. Mlati Ochieng
3. Ahmed Haji
4. Allan Omollo
5. Stella Ndegwa

Presented to Name of TM

Moringa School

Data Science Department: DSF-PT05

**Project Overview**

***Problem Statement***

Arabuko-Sokoke Forest (ASF) is a biodiversity hotspot, the most significant remaining stretch of dry coastal forest and UNESCO’s Man and Biosphere Reserve (MAB). The forest continues to face threats from illegal poaching and logging emanating from the over 200,000 persons living at the forest boundary and demand for forest products from Malindi, Watamu, Kilifi and Mombasa. Developing effective strategies to combat illegal utilization of forest resources requires a deep understanding of their patterns, drivers and potential interventions. With historical law-enforcement and biodiversity monitoring data, we will use machine learning to develop predictive models to forecast future illegal activities in the forest reserve and support proactive conservation efforts.

***General objectives***

Provide actionable insights to guide conservation interventions and law enforcement strategies.

***Specific Objectives***

1. Based on historical data, develop predictive models to forecast illegal logging and bushmeat poaching activity.
2. Identify key environmental, temporal and spatial factors influencing illegal activities.
3. Evaluate the effectiveness of different machine learning algorithms in predicting and understanding illegal logging and poaching patterns.

***Data Collection***

Retrieve historical records detailing incidents of illegal logging and bushmeat poaching from diverse sources, encompassing governmental agencies, non-governmental organizations (NGOs), research institutions, and community accounts.

Compile environmental datasets covering aspects such as land cover, vegetation density, proximity to roads, and locations of protected areas.

Prioritize data quality and consistency by rigorously preprocessing and cleansing the datasets.

***Methodology***

Explore various machine learning algorithms, including time series forecasting models (e.g., ARIMA, LSTM), ensemble methods (e.g., Random Forest, Gradient Boosting), and deep learning architectures (e.g., neural networks).

Split the data into training, validation, and testing sets to evaluate model performance and generalization.

Conduct feature engineering to select relevant predictors and transform variables if necessary.

Apply techniques such as cross-validation and hyperparameter tuning to optimize model performance.

Utilize spatial analysis tools to incorporate geographical information and identify spatial patterns of illegal activities.

**Analysis and Interpretation:**

Assess the predictive accuracy of developed models using appropriate evaluation metrics (e.g., RMSE, MAE, accuracy).

Interpret model results to understand the relative importance of different predictors and their relationships with illegal logging and poaching activity.

Conduct sensitivity analysis to evaluate the robustness of the models to changes in input variables and parameter settings.

Visualize model outputs and spatial patterns of predicted illegal activities to communicate findings effectively to stakeholders.

**Implementation, Deployment and Output:**

Develop a user-friendly interface or dashboard to allow stakeholders (e.g., conservation practitioners and law enforcement agencies) to interact with the predictive models and access actionable insights.

Provide training and capacity-building workshops to relevant stakeholders using machine learning tools for conservation planning and enforcement.

Continuously update and refine the predictive models based on new data and user feedback to improve their accuracy and relevance over time.