

+254 725-022-596
koome09@gmail.com

Koome Derrick

Geospatial Data Scientist

linkedin.com/koomederrick
github.com/koome-dev
[My Portfolio website](#)

EDUCATION

The University of Nairobi, Kenya.

Bsc. Geospatial Engineering

August 2013

Relevant Courses: Remote Sensing, Geographical Information Systems (GIS), Digital Cartography

SKILLS

Programming Languages: SQL, Python

Machine learning: familiar with the Scikit-Learn library and algorithms such as logistic regression, support vector machines, decision trees, random forests, k-fold cross validation, hyper-parameter tuning etc.

Software's: Pycharm, Jupyter notebooks, QGIS, ArcGIS, Visual Studio Code

Database: MySQL

Geospatial Analysis: familiar with geopython libraries such as Geopandas, Geopy, Plotly, Folium, Opencage and can perform spatial operations and geocoding on data to generate visuals like Choropleth, Geographical scatter plot, Geographical Heatmap, Markers and Marker Clusters. Currently learning Google Earth Engine for Machine Learning , Change Detection and Land Use Land Cover (LULC) Mapping.

Core Competencies: Writing – Co-author of an industry book *“Project Design for Geomatic Surveyors and Engineers”* which is available for sale on **Amazon**.

PROJECT EXPERIENCE

January 2024 – February 2024

Machine learning project with deployment: American President OpenCV Classification

This machine-learning image classification project makes use of Python's opencv and wavelet transform libraries to detect and classify images of 9 US presidents. The workflow involved the following:

- **Data Collection:** Went on Google and batch-downloaded about 150 photos of each of the 9 presidents. Used the Fatkun tool to batch-download them.
- **Data Cleaning:** Utilized haarcascade eye and frontface algorithms to detect the face and two eyes. This was such that once the face and two eyes have been detected, the entire face would be cropped and stored in a folder. Wrote Python code that automated the process of visiting folders, carrying out haarcascade and storing the cropped images in a new folder.
- **Feature Engineering:** Used Python's wavelet transform library to process the images by a Fourier transform process and obtain mapped images of the cropped photos. Both photos(Cropped + wavelet transform) are then converted into a Numpy array and vertically stacked to improve face and eye recognition.
- **Model building:** Using Python's scikit-learn Library imported modules such as StandardScaler, Pipeline, SVC, train_test_split to carry out an initial build.
- **Hyper-parameter tuning:** Used GridSearchCV paired with a model parameter dictionary to carry out a five-fold cross-validation to get which model performed best. Compared Support Vector Machines, Random Forest and Logistic regression techniques. Random Forest had a score of 40%, SVC 75%, LR 72% with the training dataset. SVC scored 83% and LR 86% with the test dataset. Since the average was a tie between SVC and LR, settled on SVC since it came out on top after a five-fold cross-validation.
- **Model Saving:** Saved the model (pickle file) and name dictionary using Joblib and JSON modules.
- **Model Deployment:** Used Streamlit Python library to build the UI of the final web app. Wrote Python code surrounding the saved model that will allow users to upload an image of any of the 9 US presidents and get a

report of who the president is together with a class probability report. Finally, I deployed the model on Heroku with this link: <https://american-president-classifier-a8dad2a561a0.herokuapp.com/>

March 2024

Global warming Geospatial Analysis

Analyzed global average temperatures of the world with data spanning more than 200 years (1750-2020) and developed beautiful visualizations using line charts, choropleth maps and geographical heat maps. Was able to extract the relevant information using my python and pandas library skills. Was able to create visualizations to show that average global temperatures have been increasing with the rise in industrialization.

April 2024

Covid 19 Geospatial Analysis

Analyzed Global Covid 19 data acquired during the pandemic. The data included confirmed, recovered and death cases from the beginning of the pandemic to February 2022. I was able to create animated choropleth maps that show the rate of spread of the disease as well as the recovery rate. Also created animated geographical scatter plots that showcased the same information with a different visualization. I also explored the use of markers and marker clusters to demonstrate the cumulative covid 19 deaths per country.

WORK EXPERIENCE

April 2014-Current

Various: Geospatial Surveyor

I worked in various capacities as a Surveyor, two years for a busy geospatial firm and then the next eight as an independent surveyor. I worked on projects all across East Africa (Kenya, Rwanda, Somalia and South Sudan) and even became a licensed drone pilot and co-authored an industry book as indicated earlier. After about nine years in the industry, I was enamored by the tech world and spent the last one year and a half exploring where I would fit. I explored UX design and got some certifications from Google but I found my interests and efforts gravitating towards Python and then subsequently into Data Science and Machine Learning.

CERTIFICATES

- Kaggle: Geospatial Analysis
- Udemy: Spatial Analysis and Geospatial Data Science in Python
- Udemy: Google Earth Engine for Machine Learning and Change Detection
- Udemy: Complete Tensorflow 2 and Keras Deep Learning Bootcamp(Ongoing)

INTERESTS AND HOBBIES

- **Favorite books:** Virtuous Leadership (Alexander Havard), Zero to One (Peter Thiel), Linchpin (Seth Godin), Atomic Habits (James Clear).
- **Natural abilities:** I write well. I wrote dozens of articles in my previous industry which got published in leading industry magazines at the world stage.