


Elvis David

Data | Machine Learning Engineer,

I am a Data | Machine Learning Engineer with a strong mathematical background and experience in software development, creating data pipelines, feeding into deep-learning and machine-learning models, and deploying and maintaining solutions in production. I enjoy using ML, Computer Vision, and NLP models to create algorithms that utilize data from an ETL pipeline to provide valuable insights.

komboelvis08@gmail.com 

+254716929749 

Nairobi, Kenya 

linkedin.com/in/david-elvis 

github.com/Davidelvis 

WORK EXPERIENCE

Machine Learning Engineer Kunumi

03/2021 - Present

Belo Horizonte - Brazil

Achievements/Tasks

- Developing a process-oriented data platform that behaves like a traditional database and automatic prediction engine that is to help businesses save billions by answering questions from data.
- Helped a client company to save 10% of its expenses using reinforcement learning to predict, allocate and direct tug boats used for towing cargo ships.
- Improved the delivery of projects by internal data scientists 3x faster by building template pipelines that were fine-tuned to specific projects.
- **Tech stack: Python, Git, Stats models, Sklearn, Snowflake, Tensorflow, Kedro, Pandas, Numpy**

Lead Machine Learning Engineer Omdena

11/2021 - 08/2022

San Diego CA - USA

Achievements/Tasks

- Creating a hybrid model of computer vision and NLP techniques to perform near-accurate data extraction.
- Capturing KPI value indicators and extraction of information using a NER-based approach from spacy.
- Led a team in creating the Dagshub plus DVC data pipelines and also designed the deployment and packaging of the project.
- **Tech stack : Python, DVC , Git , NLP, Apache Airflow, Kafka, AWS , Prefect**

Machine Learning Engineer Omdena

11/2020 - 06/2021

New York NY - USA

Achievements/Tasks

- Built a Knowledge Graph of land ownership in Kenya, which boosted the efforts to restore degraded land.
- Built a Named Entity Recognition model (Spacy, Transformer), for entity identification from gazettes and court documents.
- Geo-tagged network data which made it easily accessible to social activists, and journalists.
- **Tech stack : Python, Neo4j, NLP, Git, Apache spark, Kafka**

PERSONAL PROJECTS

Data Augmentation using M-BIGAN in Network Intrusion Detection Systems

- The goal of the project is to deal with the class imbalance problem of cyber attacks in network intrusion detection systems.
- Built M-BIGAN model which generates more data for the minority classes. The model has multiple generators which help to cover the whole data space and by this, it attains more accuracy than the standard BIGAN.

SKILLS

Python

GCP

TensorFlow

Docker

Apache Kafka

Apache Spark

Kubeflow

Git

Bash

AWS(Redshift, Glue, RDS, Athena, S3, Kinesis, Lambda)

DVC

NLP

ETL Pipelines

Apache Airflow

Postgresql

Snowflake

SQL

ACHIEVEMENTS

Google Cloud Certification

Completed Google Cloud training to have fundamental understanding of IT services and their uses in Google Cloud and demonstrated cloud fluency and foundational Google cloud knowledge

COURSERA, Machine Learning Engineering for Production(MLOps)

Attained a certificate in MLOps course offered on coursera by DeepLearning.AI, The scope of the course included Project Scoping and Design, workflow orchestration, maintaining and monitoring a continuously operating production system

EDUCATION

B.Sc in Pure and Applied Mathematics Jomo Kenyatta University of Agriculture and Technology (09/2017 - 06/2020)

LANGUAGES

English

Full Professional Proficiency

Swahili

Native or Bilingual Proficiency