Adah Kibet

A Data Scientist and Machine Learning Engineer ready for remote work or relocation with experience working in Bio Tech, AI, Ed Tech and Software companies using Python and R. I have a keen interest in research and combining it with engineering implementations to develop products and create informative, decisive and sustainable production workflows.

adakibet@gmail.com

+254 745 311-001

MOMBASA, Kenya

linkedin.com/in/ada-kibet-877338189

github.com/ada-k

medium.com/@adakibet

SKILLS

Python | R | Theano | TensorFlow | OpenCV | ScikitLearn | PyTorch | NLTK | Bayesian Inf. | Heroku | Git | NLP | QGIS | EC2 | SQL | Spark | SciPy | ETL | Conversational AI | S3 | Matlab | Web Scraping | Tableau | Power BI | Hadoop | Kubeflow

PROJECTS

Tweets Topic Modeling & Sentiment Analysis

Recommendation System

Churn Prediction

Music Genre Classification

Biometric Attendance

WORK EXPERIENCE

Data Scientist
Pivot Bio.

02/2021 - Present   *Berkeley, USA. (Remote).*
Bio Technology Company.

Achievements/Tasks

- Automated detection of pass/fail and harvest time for microbial strains given specific conditions by building a CNN time-series classification model using keras and TensorFlow with a 91% accuracy.
- Quantified product performance through analysis of data from 8 field trials by fitting a Bayesian mixed-effects regression model using lme4 to quantify effect sizes of different treatment groups and calculated the Minimum Detectable Difference for each treatment and field.
- Led the Data Engineering support team.
- Disqualified the validity of satellite imagery in detecting product effect by bagging mixed effect regressors for 30 trials data after correcting for spatial auto-correlation using Gaussian Process and Spatial Error models. Generated field maps using ggplot2.
- Built python libraries to: generate reports on the status of data in the Warehouse and set satellite imagery streaming workflows through automated API calls from hubs like Planet and Descartes Labs.
- Designed, prototyped, implemented, evaluated, deployed a Lidar Toolkit app in Python and Bash.

Machine Learning Researcher
Traits AI.

01/2021 - 08/2021   *San Diego, USA (Remote).*

Achievements/Tasks

- Researched, built, trained, tested, optimised and deployed the FastSpeech2 Text to Speech Conversational AI model using RNNS and GANS in PyTorch on remotely hosted supercomputers.
- Researched, built, trained, tested and optimised the Synthesize Obama Speech to Video model in PyTorch and Scikit Learn on AWS EC2. Used GitHub for collaboration.
- Optimised loss functions through hyperparameters tuning and vocoders change with the aid of a TensorBoard and MFA.

Data Science Intern
10 Academy

06/2020 - 10/2020   *Remote*

Achievements/Tasks

- Designed, built, tested & deployed (using Heroku) a web app that predicts Pharmaceutical Sales across multiple stores & analyses existing data. The Random Forest Regressor model had an MSE of 0.01876.
- Performed change point analysis using Theano and Pymc3 to quantify the impact of the Kenyan government policy interventions to slow the spread of COVID-19.
- Built a Covid-19 tableau dashboard after creating an ETL pipeline using awswrangler, S3, SQL & python.
- Performed sequential and classical AB Testing performance evaluation on an Ad campaign which showed a 1.2% lift in brand awareness higher than the minimum detectable change set.

Software Engineering Intern
Swahili Box Hub.

04/2019 - 10/2019   *Mombasa, Kenya.*

Achievements/Tasks

- Built a Marine ecosystem management web app using Django, JavaScript and PostgreSQL db.

EDUCATION

B.Sc Mathematics and Computer Science.
Technical University of Mombasa.

09/2016 - 02/2021   *2nd Class Honors, Upper Division.*

Certification, Spatial Analysis and Satellite Imagery in a GIS.
University of Toronto (Coursera 2021)