RUTH NDUTA

Phone: +254713297784   **E-mail:** ndutaruthmuriithi@gmail.com

DATA SCIENTIST

Innovative and result-driven Data Scientist with 2 years experience in Data Analysis, Data Visualization, Business Intelligence and Machine Learning experience. Am passionate about in learning and adopting different analysis tools to expand my proficiency and professional development. Having worked in diverse data science projects, I possess adequate competence in Python, R, SQL, PowerBI, Excel, Qlik and Tableau to excellently perform assigned tasks and responsibilities. Given the job, l will be an excellent asset to your organization which will gain from my invaluable skill and hands-on technical knowledge.

CORE SKILLS

- Hands-on experience with programming using Python (NumPy, Pandas, Scikit Learn), SQL, R.
- Understanding of Data Science concepts including but not limited to data cleaning, feature selection, cross-validation, parameter tuning, ensemble and stacking algorithms, market basket analysis etc.
- Ability to perform Statistics and Probability tasks including but not limited to hypothesis formulation and testing, sampling techniques, and accuracy interpretation.
- Planning, Design, and Implementation of projects using the agile methodology and design thinking.
- Skills of deep learning, computer vision, cloud solutions (AWS and Azure) as well as the soft skills of leadership, teamwork, and ethics critical for effective management.
- Application of business concepts in a constantly evolving environment, and knowledge sharing through teamwork.
- Skilled in stakeholder management, communicating clearly and persuasively, thinking on my feet even when under pressure.

PROFESSIONAL CERTIFICATION

- 2022:   DLI Platform Course for Instructors.
- 2021:   AWS Machine Learning Foundations.
- 2021:   Browser-based Models with TensorFlow.js.

EDUCATION

- 2021 – To Date   Masters of Big Data Technology.
- 2013 - 2018   BSc. Electrical and Telecommunication Engineering.

2022 (January) – To Date; **Data Analyst, Fidelity Shield Insurance Company Ltd**.

- Preparing data from the central repository to create daily, weekly, monthly, and quarterly reports which involved creating and analyzing the daily and weekly data for Know Your Customer (KYC) compliance reports.
- Enhanced the process of viewing reports effectively by creating PowerBI analytics dashboard for tracking and monitoring daily, weekly and monthly performance for the Underwriting and Claims departments while optimizing data capture to enable comprehensive report retrieval from the system database.
- Implemented complete setup of PowerBI Service and access structure for the team to continuously monitor, analyze, and report daily and cumulative premium production across branches to find trends and patterns in production data.
- Collaborate with supervisors and operation manager to enhance supportability and identify performance report bottlenecks while curating new reporting reports and dashboards templates.

June 2021 – Dec 2020; **Computer Vision Intern, Positive Red Communications**.

- Setting up and configuring the Nvidia jetson Nano and downloading relevant libraries in containers computer Vision in order to capture workers and product images using the jetson nano and store them for processing in machine learning algorithms
- Image annotation and augmentation using Computer Vision Annotation Tool and storing labels and annotations in the desired format that can later be fed into the model in a version that can be understood and trained on to improve detection accuracy and minimize bounding errors and misclassification.
- Design and build an object detection system for casual laborers at the industry to detect, track and count the number of products moving out of the warehouse during loading onto trucks for transport. This aimed at reducing fraudulent cases in the manufacturing plant
- Incorporate tracking and counting of casual workers bi-directionally to either add or subtract forward-moving or returning goods into the warehouse. This involved immense open-ended research to improve the quality of collected data including but not limited to semi-supervised learning and fine-tuning the machine learning model with human movement.

Sept 2018 – June 2020; **Telecommunications Engineer, Innovis Telco Ltd.**

- Technical support to mobile network operator Safaricom in Western Kenya. Main tasks comprised hardware expansion and software upgrade of the mobile packet core network elements as well as operations support. Having done this, the operator was able to significantly improve their customers' user experience especially in the interior regions of western Kenya.
- Supervised installation and optimization of about 200 network sites for 4G and 3G upgrade. The assignment required performing standard operational procedures and checks; investigating and diagnosing faults reported in the data network; collection system fault data and logs; I accomplished significant improvement in incident management turnaround time as well as greater operator confidence out of the knowledge transfer efforts.

- Managed quality assurance program, including on-site evaluations, internal audits, and customer surveys which resulted in improved network performance indicators that enhanced the data network customer experience. Subsequently, major bottlenecks in the data network speed and performance were resolved.
- Project assignments involved planning and implementing network expansion for both 2G & 3G network elements enabling Safaricom to enhance their data network capacity.

DATA SCIENCE PROJECTS

1.   Building Transformer-Based Natural Language Processing Applications: Certificate
- Practical hands-on guide to the NVIDIA NeMo API and exercises to build a text classification task and a named entity recognition task using BERT-based language models

2.   Web3 Gender Hackathon April 2022:
- An introduction to Web3 using Polygon and Harmony protocols to build scalable blockchain-based applications.
- Built a scalable web-based blockchain application to help reduce maternal mortality rates in low-income households.
- The dapp is accessible through USSD to allow for inclusivity.

3.   UmojaHack Africa 2022: Monthly Insurance Claim Prediction Challenge: Certificate
- The objective of this challenge is to create a machine learning model to predict how much a client will claim from Zimnat per month for a whole year.
- Used Catboost Regression algorithm for prediction.
- Performed feature engineering and hyperparameter tuning to improve model performance.

4.   AFD-Gender-Based-Violence-Dataset-Collection-Challenge May 2021: GitHub
- The objective of this project was to collect data on the internet related to gender-based violence cases among women in Africa
- Used Tweepy to scrape Twitter hashtags related to the case.
- Main programming language - Python.

5.   Car-insurance-prediction April 2021: GitHub
- The objective was to build a machine learning model to predict if these customers will buy the car insurance or not.
- Used LightGBM classifier and Random Forest classifier to make predictions.
- Achieved an accuracy of 70.56% using the LightGBM classifier that performed best.
- Main programming language - Python.

6.   Financial Inclusion Analysis October 2020: GitHub
- The main objective was to predict which individuals are most likely to have or use a bank account.
- Used Random Forest classifier to make predictions.
- Achieved an accuracy of 85.84%.
- Main programming language - Python.

AI COMMUNITY CONTRIBUTION

- Zindi - Country ambassador Kenya. Leading hackathons and promoting Zindi events in local universities.
- Deep Learning Indaba: Eastern Africa Representative. Organize IndabaX among AI communities, budget, and follow-up expenditure reports, coordinate communication between Indaba and the East Africa region.
- Masakhane: community contributor in NLP research in Africa.
- Hackmakers Regional Ambassador: organizing the United Nations Innovation Day Hackathon.

_____

REFEREES:

1. Mr. Patrick Mukunzi
   Reinsurance Manager,
   Fidelity Shield Insurance Company Ltd,
   Email : pmukunzi@fidelityshield.com

2. Ms. Rose Delilah
   Technical Mentor,
   Moringa School.
   Email : rose.delilah@moringaschool.com

3. Ms. Mercy Cheruiyot
   Data Scientist,
   KPMG.
   Email : cherono.mj@gmail.com