PATIENCE SIMULI SIMIYU

0746868070 | patiencesimuli19@gmail.com | Nairobi, Kenya | LinkedIn

SUMMARY

- Multi-domain Machine Learning engineer with **3+ years' experience** in the **Telecommunications**, **Finance**, **retail**, and manufacturing sectors.
- A lifelong love and exceptional ability for **research, innovation and analytics** led me to extend my skills beyond RAN Optimization Engineering/Telecommunications Engineering and explore big data analytics and Machine Learning Engineering early on.
- Proficiency in **Natural Language Processing**, **Deep Learning**, **Computer Vision**, **Time series and regression analysis, Reinforcement Learning**, data mining large data sets, and Python backend web development using **Django**, **Flask**, and **Fast API** in conjunction with **MongoDB**.
- Extensive data modelling and data analysis experience with exposure to the **Hadoop ecosystem**, AWS, Huawei, and Private **Cloud deployments**, Python, and **backend SQL**.
- Expertise in **ETL operations** and tools and report creation using **PowerBI, Tableau**, and DataLink.
- Professional experience installing Hadoop in AWS and extracting data from traditional database to **HDFS**.
- Exceptional people skills and ability to lead and collaborate with multi-domain teams.
- Passionate in leveraging my technical skills to contribute to community development initiatives, especially those aimed at **protecting and improving the lives of children**.

SKILLS & TOOLS

- Machine Learning: Python, Sci-Kit Learn, TensorFlow, PyTorch, PyCaffe, Keras, Natural Language Processing (NLP), Computer Vision, Regression and Time series analysis, Feature Engineering, Git version control, Agile project management, and MATLAB.
- Analytics and Visualization: Python libraries, Machine Learning, Deep Learning, GIS Data analysis, Telecommunications Big Data Analysis, Financial data analysis, IBM Cognos, R Studio, R-Programming, Tableau, Advanced Excel VBA, Macro, Microsoft Power BI.
- Data Management: Kimball-DW/BI Life cycle methodology, Erwin, Microsoft Visio, MySQL, Mongo DB, Oracle 11g, and Oracle 12c.
- Programming and Scripting Languages: Python, Java, SQL, UNIX shell scripting, C.
- Domain: Telecommunications, GIS, Finance.
- Operating systems: Ubuntu, Windows, UNIX.
- Hadoop Tools & Concepts: HDFS, Map Reduce, CDH5 Hadoop Cluster, Hive, Spark, Impala, Hbase, Flume, Kafka.
- RAN Engineering: Nokia NetAct, Huawei iMaster (PRS, U2020), Genex Cloud, Probe & Assistant, U-Net, Naster, WMG (Wireless Fault Management Analysis by Huawei), Nokia NSP, Nemo Outdoor, TEMS Investigation, and Atoll.

EDUCATION

AWS Certified Solutions Architect Associate **July 2023**
Amazon Web Services

Huawei Certified IT Associate – Cloud Services **July 2024**
Huawei Technologies

Bsc. Telecommunication and Information Engineering **December 2023**
Jomo Kenyatta University of Agriculture and Technology

AWS Machine Learning **July 2021**
Amazon Web Services

Kenya Certificate of Secondary Education (KCSE) **November 2017**
Moi Girls High School Eldoret

EXPERIENCE

Data Scientist | RAN Optimization Engineer **February 2024 - Date**
Huawei Technologies Kenya Limited

- Bridge between the RAN Optimization team and SOC Big data team.
- End customer value management: Built predictive models using machine learning algorithms to perform predictive analytics.
- Implemented machine learning algorithms using MLlib and GraphX APIs for graph-parallel computation.
- Sentiment analysis for client's broadband and GPON solutions.
- Utilizing ArcPy module for Python geoprocessing to profile end users and improve RF services.
- Data pipelines design and migrations for telecom billing, financial, and UE Measurement Reports data.
- Gained hands-on experience with the Hadoop ecosystem, managing data distribution, storage, and processing in a Hadoop cluster deployed in the cloud.
- Worked with JSON scripts, MongoDB, and Unix environment to perform non-SQL data clean-up and grouping, and generated analysis reports for insights.
- Customer value GIS-based analysis and management: Using QGIS and ESRI ArcGIS to create customer profiling polygons, including overall value and services present, then upselling new services.
- Developed streaming applications using Apache Spark, writing APIs in Scala, Python, and Java to process real-time data and deliver scalable data solutions.
- Strong understanding of Kerberos authentication principles and experience working with Kerberos servers to manage secure access to network resources.

- RF KPI dashboard creation, including visualization charts and reports, using Tableau, DataLink, and DataCube.
- Led and coordinated customer interactions while handling customer projects.
- All other RAN Optimization Engineering tasks.

Environment: Hadoop Ecosystem, HDFS, Hive, HBase, SQL, Python, Java, Tableau

Machine Learning & Electronics Consultant **October 2023 – June 2024**
Naivas Supermarket Limited (Via the Industry Meets Talent program)

- Conceptualized and developed a smart Machine Learning Powered IoT weighing scale. Used machine learning algorithms to empower the scale with advanced inventory management, auto-replenishment, and sales forecasting functionalities.

- Machine Learning Algorithms: Random Forest, Market Basket Analysis, and Isolation Forest.

Machine Learning Research Assistant **May 2021 – June 2024**
Jomo Kenyatta University of Agriculture and Technology

- **Scalable Reinforcement Learning [1]-[3]:** Combining Federated Learning with Hierarchical Multi-Agent Reinforcement Learning (H-MARL) to enhance the scalability of RL models in drone-based identification of underground water sources and similar applications that require the processing of diverse and extensive datasets over large geographic areas.
- **Smart Grid Deep Learning [4]-[6]:** Designed and executed a robust Deep Learning solution leveraging artificial neural networks (ANN) and a Decision Tree Classifier to swiftly identify three-phase faults, locate their positions, and classify fault types within the physical power grid. This research and development initiative aimed to solve Kenya's power outage problem. Streamlining fault location and classification will ultimately reduce downtime and improve operational efficiency.
- **Computer Vision (CNN) + Depth sensing [7], [8]:** Combining CNN and depth sensing in smart glasses using LiDAR and Stereo cameras to improve real-time 3D object mapping for navigation and avoiding obstacles.

Data Scientist | RAN Optimization Engineer **Jan 2023 – Jan 2024**
Jamii Telecommunications Limited

- NLP: Hive query and Python – based customer feedback and social media sentiment analysis.
- Machine Learning customer data profiling based on UE Measurement reports, location, and value.

- Hands-on experience with Cloudera Manager and Administrator for efficient data management, operations, and system optimization.
- Experienced with Kimball DW/BI Data ware house Life cycle.
- Real time data storage and processing with Apache Flume and Kafka.
- Data model creation using Erwin.
- Customer value GIS-based analysis and management: ESRI ArcGIS to create customer profiling polygons, including overall value and services present, then upselling new services
- Designing and optimizing queries, developing data universes, and crafting tailored reports to meet specific requirements. filtering, grouping, and transforming data to provide actionable insights
- RF KPI dashboard creation, including visualization charts and reports, using PowerBI and Tableau
- Report creation to present actionable insights using SAP Webi and Crystal Reports
- All other RAN Engineering tasks.

PERSONAL ACHIEVEMENTS

- Developed an RF interference mitigation solution including detection, isolation, and mitigation using advanced algorithms, interference coordination, and power control techniques.
- Built an ML-based cybersecurity threat hunting simulation system with behavioural analysis, baseline profile pattern recognition, and correlation and contextualization.

SKILLS

- Machine Learning and Deep Learning
- Radio Network Planning and Optimization
- SparkSQL
- Hadoop
- LAN and WAN hardware/software
- Routing Protocols
- Programming Skills: C++, C, Python, Java
- AWS and Azure Cloud computing
- Data Centre Infrastructure Setup and Maintenance
- Data Science and Artificial Intelligence
- Communication and interpersonal skills
- Analytical and problem-solving skills
- Organizational and time management skills

SELECTED PULICATIONS

[1] Hu, F., Deng, Y. and Aghvami, A.H., 2021. Scalable multi-agent reinforcement learning algorithm for wireless networks. *arXiv preprint arXiv:2108.00506*.

[2] Gogineni, K., Wei, P., Lan, T. and Venkataramani, G., 2023. Scalability Bottlenecks in Multi-Agent Reinforcement Learning Systems. *arXiv preprint arXiv:2302.05007*.

[3] Pimentel, L., Paleja, R., Wang, Z., Seraj, E., Pagan, J. and Gombolay, M., 2022. *Scaling Multi-Agent Reinforcement Learning via State Upsampling* (No. SAND2022-8384C). Sandia National Lab.(SNL-NM), Albuquerque, NM (United States).

[4] Hatziargyriou, N., 2022. Frontiers in Smart Grids. *Frontiers In Smart Grids*, *1*, p.986178.

[5] Dagle, J.E., 2012, January. Achieving resiliency by eliminating common mode failures in the smart grid. In *2012 IEEE PES Innovative Smart Grid Technologies (ISGT)* (pp. 1-2). IEEE.

[6] Gaouda, A.M., 2013, November. Adaptive Partial Discharge monitoring system for future smart grids. In *IECON 2013-39th Annual Conference of the IEEE Industrial Electronics Society* (pp. 4982-4987). IEEE.

[7] Mhaske, A., Nikam, A., Bhandari, S. and Mantri, S., 2020, November. Smart glasses to assist monocular vision to estimate depth. In *2020 IEEE International Conference for Innovation in Technology (INOCON)* (pp. 1-7). IEEE.

[8] Wang, W. and Neumann, U., 2018. Depth-aware cnn for rgb-d segmentation. In *Proceedings of the European conference on computer vision (ECCV)* (pp. 135-150).