Celdrick Ndze Kuta

Data Scientist

(309) 989-0419
Lawrenceville, GA
kutaceldrick880@gmail.com

github.com/kuta-ndze
linkedin.com/in/kuta-n-celdrick-b808ba169

EDUCATION

MSc. in Applied Statistics,	Illinois State University	**May 2021**
MSc. in Computer Science,	University of Yaoundé	**Jul 2016**
BSc. in Mathematics | Computer Science,	University of Buea	**Jul 2010**

SKILLS

Tools and Languages	Python, R, Git, MATLAB, Dash, SQL, Oracle, SAS, MySQL Developer, Azure Databricks, AWS Console, IBM Cloud, AWS SageMaker, AWS Kinesis, AWS Glue, AWS EMR, CSS, HTML5, JavaScript, LaTeX, MarkDown.
Soft Skills	Machine Learning, Deep Learning, CNN, NLP, TCN, ETL, GAN, TensorFlow, Keras API, Random Forest, LDA/QDA, XGBOOST, Power BI, Tableau, AWS QuickSight, Hyperparater Tuning, Model deployment.
Others	Communication (English & French), Hadoop, Spark, VBA, Database design, C, C++, R Shiny, SDLC.

TECHNICAL EXPERIENCE

Data Scientist   **Mar 2020 — Present**
Elevated Technologies   *Houston, TX*

- Developed model to predict insurance claim severity on client's data set with Random Forest .
- Performed Data Wrangling, Exploratory data analysis, outlier treatment, missing values,variable encoding.
- Predictive analysis of client's data set for preparation of offers and other campaigns for selling financial products.
- Performed Dimensionality reductions with PCA, Class imbalance.
- Hyper-parameter tuning with Random Search.
- Data Visualization AWS QuickSight

Data Scientist   **Apr 2019 — Feb 2020**
Optimum   *Houston, TX*

- Predictive maintenance models to develop defect detection, classification and localization models.
- Data visualization and pre-processing with Python Pandas, Matplotlib and Seaborn .
- Deep Learning TensorFlow Model development with ResNets, ResUnets, CNN.

Data Scientist   **Jul 2018 — Mar 2019**
Hirewell   *Chicago, IL*

- Web scrapping text/unstructured data from eCommerce website Home Chef with python beautiful Soup.
- Application of NLP to perform Sentiment analysis on considerably large data set.
- Feature Extraction and Modelling with Naïve Bayes classifiers model, Logistic Regression using scikit-Learn.

Data Analyst   **Nov 2016 — Jun 2018**
Panus Software   *Buea, Cameroon*

- Coordinated detailed industry analysis, research, drafted reports and developed analytical insights for Banking Clients to optimize marketing strategy by performing customer segmentation.
- Apply clustering ML technique K-means algorithm and auto-encoders, PCA.
- Created dashboards that identified KPIs, defined new columns.
- Involved in creating database objects like tables, views, procedures, triggers, and functions using T-SQL to provide definition, structure and to maintain data efficiently.

RECENT PROJECTS

Medical Insurance Premium Prediction   **Jan 2022**

- Train a Linear learner, multilayer perceptron Networks in AWS SageMaker.
- Deploy Model and invoke endpoint in SageMaker.

Disease detection Automation   **Jan 2022**

- Application of CNN architectures, transfer Learning and ResNets.
- Model evaluation and deployment.

ACHIEVEMENTS

IBM Data Science Professional Certificate	**Jan 2021**
Power BI	**Jan 2021**
AWS SAA	**Jun 2021**
Linux DBA	**Aug 2021**
Society of Actuaries Exam P/1	**May 2019**