OLUSEGUN STEPHEN OMOTUNDE

(470) 435-1887| omotundejunior@gmail.com| Marietta, GA | website | tableau| https://www.linkedin.com/in/olus/ | medium

EDUCATIONAL QUALIFICATION

Bowling Green State University, *Bowling Green, Ohio*   Dec 2022
Master of Science in Applied Statistics with a Business Analytics Specialization

- **Relevant Coursework**: Data Science Programming, Data Mining, Database Management, Probability Theory, Exploratory Data Analysis, Linear and Integer Programming, Regression Analysis, Probability Models for Decision Making

Weclouddata, *Ontario, Toronto, Canada*   June 2022
Diploma of Education in Applied Data Science & Big Data

- **Technologies Covered**: Languages (Python, SQL), Cloud Platform (AWS), Big Data (PySpark, Databricks), Machine Learning (Linear & Logistic Regression, SVM, Decision Trees, Random Forests, Neural Networks)

University of Ilorin, *Kwara State, Nigeria*   Oct 2019
Bachelor of Science in Finance

WORK EXPERIENCE

DreamChaseTechnology, Brooklyn, New York   April 2023 – Present
Data Scientist

- Developed a **chatbot** utilizing **google Med-PaLm large language models (LLM) with Langchain and RAG**, aimed at improving the customer search experience for a healthcare provider client. **Utilized Vertex AI** and **Kubernetes** on **gcp for deployment.**
- Designed and implemented efficient **ETL processes** with **PySpark** and **GCP Vertex AI** to transform and load data into the **Google Cluster Pod.**
- Created a **vector database** to store vector embeddings of prompts, enhancing search accuracy time.
- Conducted **text parsing** and generated unstructured data for all attributes in the dataset.
- Implemented **semantic and metadata search** capabilities by enabling self-query retrieval based on metadata descriptions.
- Conducted regular knowledge-sharing sessions with cross-functional teams and stakeholders, employing presentations and PowerPoint to enhance the team understanding of advanced analytical techniques, contributing to a collaborative and informed work environment.

Bowling Green State University, Bowling Green, Ohio   Aug 2021 – Dec 2022
Graduate Teaching Assistant & Data Analyst

- Automated the collection of financial data, including YTD total returns for a diverse set of assigned stock tickers (VZ, JPM, VGT, VHT, VOX, AAPL) from sources such as Finra, Barrons, Yahoo Finance using **advanced web scraping techniques, Python scripting, Beautiful Soup and Selenium automation.**
- Utilized Microsoft Excel for **regression analysis, calculating CAPM beta, and Jensen's alpha** on the collected data, providing nuanced interpretations of their financial implications. leveraged **advanced Excel functions and pivot tables** to perform data cleaning and analysis on faculty and alumni survey data on coverage of Sustainable Development Goals.

Nigeria Airspace Management Agency, Lagos, Nigeria   Jan 2020 – Dec 2020
Financial Data Analyst

- Applied advanced quantitative analysis and machine learning techniques to analyze payables data using **Python and SQL.** Detected trends and identified peak payment cycles, resulting in a significant reduction in overtime costs across the 36 states in Nigeria and for top vendors. This initiative led to a projected 8-10% cost savings.
- Presented custom dashboards to visualize key financial metrics, enhancing data accessibility and aiding in informed decision-making. Effectively communicated data-driven insights to management, thereby facilitating the successful implementation of cost-saving initiatives and streamlining financial processes.
- Architect of a high-performance data warehouse on AWS, reducing manual reporting efforts by 40%.

Caverton Helicopters, Lagos, Nigeria   Jan 2018 – Oct 2018
Business Analyst Intern

- Enhanced the client payment strategy through the creation of a predictive classification model using **Applied Logistic Regression**. This model accurately predicted and forecasted aircraft downtime with an 85% accuracy rate, enabling more informed client payment terms and contract structures.
- Actively contributed to the development of a comprehensive reporting dashboard that presented data- driven insights and actively monitored customer flight patterns, aircraft health, aircraft maintenance analytics, and maintenance information.

PROJECTS

CSV Data Viz Analysis Tool (skills/Tech Stack: python, langchain, openai gpt3.5 and gpt4 llms)
- Created a user-friendly application designed for visualizing structured data. This tool allows users to effortlessly create plots with accompanying summary explanations and easily embed them in a downloadable powerpoint file.

NBA Player Clustering (*Skills: R [tidyr, ggplot2, gridExtra, cowplot], Unsupervised machine learning)*
- Performed cluster analysis on NBA player statistics using data from 2020-2021 NBA season to group players into functional categories based on their on-court statistics and style of play using unsupervised clustering algorithms (K-means, hierarchical, and model-based clustering).

Implementation of an E-commerce System on AWS in an Automated way using Terraform and Ansible (skills: Terraform, Ansible, Amazon EC2, Amazon Virtual Private Cloud (VPC))
- Utilized terraform to deploy the AWS EC2 infrastructure within Amazon Virtual Private Cloud (VPC). Utilized Ansible for the configuration management and to automate the installation of PHP, MySQL, Redis and Magento.

Implementation and deployment of a Scalable Web Application (skills: AWS Elastic Beanstalk, DynamoDB, Amazon CloudFront, Amazon Edge Locations)
- Amazon Elastic Beanstalk was used to create and deploy a web application. Amazon Auto Scaling was used to monitor the web application performance. A NoSQL Amazon DynamoDB database was created to store user emails and Amazon CloudFront distributed this web applications to edge locations around the world for faster access by users and to reduce latency.

Interactive Tableau Dashboard visualizing King County home sales data to deliver insightful market overview

TECHNICAL SKILLS

- **Programming**: Python (NumPy, Pandas, Scikit-Learn), R Programming
- **Database:** MySQL, Amazon Redshift, Google Big Query
- **Cloud technology**: AWS (Lambda, S3, EC2, Sagemaker, Glue, Athena), Google Cloud Platform (Big Query, Vertex AI)
- **Orchestration Tools**: Apache Airflow, DBT
- **Big Data**: Apache Spark (PySpark), Databricks, Snowflake
- **CI/CD**: Git, Docker, Kubernetes, Jenkins, Terraform, Ansible
- **Machine Learning & Deep Learning:** Classification, Clustering, Regression, Deep Learning, Natural Language Processing, Large Language Models, Recommender System, Computer Vision.
- **Statistics:** Probability, Distribution, Hypothesis Testing, A/B Testing, Regression, Time Series Analysis
- **Visualization:** AWS Quicksight, Tableau, Python(Matplotlib, Seaborn, Plotly), R(GGPlot), MS Excel (Charts), Power BI