Samuel Affum Kyeremeh

affum3331@gmail.com | (+233) 537854001 | Accra, Ghana | linkedin.com/in/samuelkyeremeh

EDUCATION

Kwame Nkrumah University of Science and Technology   Sep 2021 – present
Bachelor of Science, Computer Science | Second Class Upper, CWA 63.17 /100   Kumasi, Ghana
Relevant Coursework:
Databases, Python, Operations Research, Statistics & Probability, Financial Accounting, Algebra, Software Engineering

PROFESSIONAL EXPERIENCE

TECHSTRIPPED AFRICA   Jan 2023 – present
Campus Lead   Kumasi, Ghana

- Onboarded and trained 30 new team members, enhancing collaboration and teamwork within the organization by implementing structured training programs and fostering a supportive environment
- Mentored junior team members in career planning, leading to successful internships and job placements by providing tailored guidance and support, ultimately improving their professional growth and confidence
- Organized webinars featuring industry experts to provide team members with insights into opportunities in technology, aimed at expanding their knowledge and networking, thus preparing them for future career advancements

PROJECTS

End-to-End Text Summarization Pipeline using Machine Learning ⧉

- Conducted feature engineering, including tokenization, stop-word removal, and text normalization using Python libraries like NLTK spaCy to improve model accuracy and efficiency by ensuring clean and consistent input data
- Developed high-quality text summarization models utilizing Hugging Face Transformers and PyTorch to generate coherent and concise summaries, significantly enhancing the user experience in content consumption
- Designed and implemented automated data ingestion, validation, and transformation pipelines with Pandas, reducing manual errors and improving data consistency for a more reliable model training process
- Deployed the text summarization model for real-time use using AWS services (Lambda and EC2) and Hugging Face's Inference API, allowing end-users to access summarization capabilities on-demand and making the tool more accessible for various applications

Credit Card Fraud Detection using Machine Learning ⧉

- Performed feature engineering and data preparation to optimize model performance, utilizing Python libraries like Pandas and NumPy to ensure high-quality data input for model training
- Built and trained various models, including Logistic Regression, KNN, Decision Trees, Random Forest, XGBoost, and SVM, leveraging scikit-learn for implementing different algorithms to compare their effectiveness in solving the problem at hand
- Applied repeated K-Folds and grouped K-Folds cross-validation techniques, along with SMOTE and ADASYN for synthetic data generation. These techniques helped address class imbalance and ensured robust validation. XGBoost, paired with Random Oversampling and StratifiedKFoldCV, delivered the best performance in oversampled cases, achieving a ROC-AUC of 0.98 and an optimal threshold of 0.0172

911 Call Data Analysis and Visualization in Python ⧉

- Analyzed 911 call data to uncover key emergency response insights, utilizing data exploration techniques to identify trends and patterns that inform emergency response strategies
- Cleaned and prepared the data using Pandas, ensuring its accuracy and reliability to eliminate inconsistencies and prepare it for in-depth analysis
- Created detailed data visualizations with Matplotlib and Seaborn, facilitating informed decision-making for stakeholders by providing clear visual representations of critical emergency response metrics

Urban Mart Food Inventory Analysis using Power BI ⧉

- Analyzed and visualized inventory data to uncover business insights.
- Cleaned and preprocessed data using PowerQuery for accuracy and reliability
- Utilized DAX for advanced data manipulation, including calculated columns and measures
- Developed comprehensive data visualizations in Power BI to support data-driven decision-making

SKILLS

Programming Languages
Python, SQL

Data Analytics
Data Visualization, Data Integrity, Data Presentation, Data Analysis, Data Collection, Data Aggregation, Power BI, Microsoft Excel

Frameworks & Tools
Pytorch, Sklearn, Machine Learning, Pandas,Seaborn, Matplotlib, Jupyter Notebooks ,GIT, AWS, PySpark