

CLEAVESTONE ADUNGO

Data Scientist/Data Analyst

Experienced Business Analyst at Turing with expertise in Generative AI and Large Language Models (LLMs), including fine-tuning using Supervised Fine-Tuning (SFT) and Reinforcement Learning with Human Feedback (RLHF). Skilled in leveraging LLMs to perform advanced data analysis, develop innovative solutions, and bridge AI capabilities with business needs. Proficient in Power BI, SQL, Python, and Excel, adept at transforming complex datasets into actionable insights, creating dynamic dashboards, and optimizing processes for strategic decision-making. Demonstrated experience in developing AI-powered applications and integrating cutting-edge technology for impactful results in business processes. Committed to continuous learning, actively pursuing self-study and hands-on projects to stay at the forefront of data and AI trends. Seeking roles as a Data Analyst, Data Scientist, or Business Analyst to apply analytical mindset, technical expertise, and innovative approach for organizational success.



Skills

Qualitative research

Quantitative research

Data analytics - Python , SQL, Excel, Power BI

Statistical Analysis

Machine Learning - scikit-learn, TensorFlow ,PyTorch

Version Control with GIT

Computer Vision

Cloud Platforms - AWS (S3, EC2, SageMaker)

NLP – Transformers, Hugging Face, spaCy



Contact

Address

50400, Nairobi Kenya

Phone

+254703457427

E-mail

cleavestone94@gmail.com

LINKEDLN

<https://www.linkedin.com/in/cleavestone-adungo-b3b7ba124/>

GITHUB

<https://github.com/cleavestone?tab=repositories>



Work History

Nov 2024
– April
2025

BUSINESS AND RESEARCH ANALYST

TURING COMPANY, Remote

- Data Development: Curate, preprocess, and structure high-quality datasets for fine-tuning LLMs, ensuring alignment with project objectives and real-world use cases.

- Prompt Engineering: Design, test, and iterate on effective prompts to optimize model outputs for clarity, relevance, and precision.
- Ideal Response Design: Define and validate gold-standard responses to ensure models align with business goals, ethical guidelines, and user expectations.
- Fine-Tuning Support: Collaborate with ML engineers to train and evaluate LLMs, improving contextual understanding and output quality.
- Quality Assurance: Conduct rigorous validation and testing of LLM outputs, refining datasets and prompts to enhance performance.
- Cross-Functional Collaboration: Work closely with data scientists, product teams, and engineers to integrate LLM capabilities into workflows.

**Sep 2023 to
Oct 2024**

CREDIT CONTROLLER

SIGNATURE HEALTHCARE LTD, Nairobi

- Managed debt collection by monitoring account receivables, making cold calls, and planning weekly visits.
- Maintained and updated payment records, reconciled accounts, and addressed finance-related queries.
- Collaborated with the sales team to streamline debt recovery efforts and improve financial reporting.
- Led corporate banking operations and ensured compliance with payment processing protocols.
- Collaborated with other departments to resolve billing errors, ensuring timely payment from clients.

**June 2021
to Sep 2023**

ACCOUNTS CLERK

SIGNATURE HEALTHCARE LTD, Nairobi

- Managed petty cash issuance, reconciliation, and replenishment.
- Verified customer orders against price lists and credit limits for approval.
- Processed payment vouchers, receipts, and fiscal transactions while ensuring accuracy.

- Compiled return on investment reports and performed revenue analysis.
- Handled supplier LPOs, invoice approvals, and cheque record management.
- Set up and updated Excel spreadsheets to track financial data.



Education

Oct 2012 to
June 2017

Bachelor of Science: Mathematics

Moi University - Eldoret

Second Class Honours Upper Division



Certifications

June 2024
to July 2024

Certificate in Foundational Generative AI - INeuron

Oct 2024 to
Nov 2024

Certificate in LLM Engineering –UDEMY



Projects

Web Scraping Project (Python)

- Developed a web scraping application using Python to extract job listings from the Brighter Monday website.
- Utilized BeautifulSoup and Requests library for web scraping, data preprocessing, and cleaning.
- Stored extracted data in a Pandas DataFrame.
- GitHub Repository:
<https://github.com/cleavestone/Data-Analytics/blob/main/SCRAPING%20BRIGHTER%20MONDAY%20JOB%20LISTINGS.ipynb>

Sales Dashboard (Power BI)

- Designed and developed an interactive sales dashboard to analyze sales performance metrics.
- Provided insights into revenue, profit margins, top-selling products, and categories.
- Utilized Power BI modeling capabilities for dynamic data visualization.
- GitHub Repository:
- https://github.com/cleavestone/Data-Analytics/blob/main/sales_dashboard.pbix

Time Series Forecasting (Amazon closing stock prediction)

- Forecasted the closing stock prices of Amazon using machine learning techniques.
- Collected and preprocessed historical stock prices data using Pandas and NumPy.
- Developed predictive models and evaluated performance using appropriate metrics.
- GitHub Repository:
- https://github.com/cleavestone/Data-Analytics/blob/main/AMAZON_STOCK.ipynb

Data Cleaning (Glassdoor data science job)

- Conducted data cleaning focused on Glassdoor data science job listings.
- Performed data preprocessing tasks such as handling missing values and removing duplicates.
- Utilized Pandas for data manipulation and cleaning, applying regex and data type conversion techniques.
- GitHub Repository:
- https://github.com/cleavestone/DataAnalytics/blob/main/Data_cleaning_glassdoor_data_science_jobs.ipynb

Ecommerce Database –SQL

In this project, I worked with an e-commerce dataset to perform data manipulation and analysis using SQL. The database consisted of tables for customers, orders, products, order items, and reviews. I created and populated these tables, ensuring data integrity and referential integrity with foreign key constraints. This project highlights my ability to work with complex data structures, perform

data analysis, and generate actionable insights through advanced SQL querying techniques.

GitHub Repository:

<https://github.com/cleavestone/Data-Analytics/blob/main/ECOMMERCE%20PROJECT.sql>

Predicting Air Quality with Time Series Analysis

Developed a comprehensive air quality forecasting model using time series analysis. Employed advanced data processing techniques and statistical models to predict air quality indices.

Implemented machine learning algorithms to analyze historical data .

GitHub:

https://github.com/cleavestone/Air_quality_Forecasting

References

KELVIN KINYUA
HUMAN RESOURCE OFFICER
SIGNATURE HEALTHCARE LTD
0741054159

JUSTINE MOKUA RATEMO
TENDERING OFFICER
SIGNATURE HEALTHCARE LTD
0713720319

ARCHANA THIAGARAJAN
PROJECT REVIEWER
TURING LIMITED
archana.t@turing.com