Abdul Wahab Basit

(233) 542-789495 | Q Personal Website | Github | LinkedIn

RESEARCH INTERESTS

Big Data Processing, Cloud Computing, Machine Learning, Large Language Model, Natural Language Processing

EDUCATION

University of Mines and Technology | BSc in Computer Science and Engineering *2019 – 2023*

Bachelor of Science in Computer Science and Engineering (Second Class Upper Hons)

Coursework: Artificial Intelligence, Programming (C++, C#, Java), Operating Systems, Embedded Systems, Information Theory, Linear Algebra, Signal and Image Processing, Database Systems, Data Structures and Algorithms

TECHNICAL SKILLS

Languages : C++, Python, Javascript, Flask and SQL
Frameworks : OpenCV, PyTorch, Keras, Pandas, Flask, Numpy, Pyspark, Apache Airflow, NodeJS
Tools : Apache Kafka, Apache Spark, AWS (S3, EMR, Redshift, Glue)

RESEARCH PUBLICATION

- Emmanuel Daanoba Sunkari, Abdullah Bello Muhammed & **Abdul Wahab Basit**, "Application of Machine Learning Algorithms to Predict Rock Types Using Geochemical Data: A Case Study from the Obuasi Gold District, Ghana" *[To be published by Springer in Chapter 8 of Gold and Mineralization in and around Ghana Book on September 15, 2025]*
- **Abdul Wahab Basit**, George Essah Yaw Okai, Abdul-Majeed Osman, Ebenezer Afful Junior, "A review on machine learning-based sentiment analysis in combating misinformation: techniques, challenges, and applications", *2024 VI. International Turkic World Congress on Science and Engineering, Azerbaijan-Baku, 2024,* pp. 6.26-6.35

EXPERIENCE

Amalitech **East Legon, Accra, Ghana**

Data Engineer April 2025 - Present

- Developed and orchestrated batch and real-time ETL workflows using Apache Airflow, AWS MWAA, and AWS Glue, automating the ingestion and transformation of data from Amazon S3 and Aurora MySQL into Amazon Redshift for analytics and reporting.
- Wrote optimized SQL queries and scripts for large-scale data cleansing, indexing, and transformations, resulting in 65% improvement in query execution time and performance.

Part-time Research Assistant January 2024 - Present

Collaborated with Assistant Professor Emmanuel Daanoba Sunkari

Application of machine learning algorithms to predict rock types using geochemical data: A case study from the Obuasi Gold District mining district, Ghana

- Collaborated on a project to predict rock types using machine learning algorithms based on geochemical data from the Obuasi Gold District in Ghana
- Implemented supervised classification algorithms, including Random Forest, Gradient Boosting, AdaBoost, and Support Vector Machine to analyze rock samples

Hydrogeochemical Appraisal Of The Co-Occurrence Of Fluoride And Nitrate In Groundwater Of Gushegu Municipality In The Northen Region, Ghana

○ Developed machine learning models to predict fluoride and nitrate concentrations in groundwater, using multiple regression techniques including XGBoost, SVM, and Stochastic Gradient Descent (SGD)

○ Evaluated model performance using Mean Squared Error (MSE) and Root Mean Square Error (RMSE), identifying XGBoost as the most accurate model for predicting water quality

Collaborated with Dr. Affum
Hypothyroid Disease Prediction Using Machine Learning

○ Collaborated on a project to predict hypothyroid disease using machine learning algorithms based on dataset of 1,000 individuals, 376 positive cases and 624 negative cases of hypothyroidism

○ Developed and evaluated machine learning models to predict hypothyroidism using classification techniques, including ANN, Random Forest, SVM, and Logistic Regression

Ayadata (Freelance)

Data Annotator & ML Data Entry Specialist November 2024 – February 2025

○ Annotated and labeled datasets for machine learning projects, ensuring high-quality training data.

○ Worked with remote teams to maintain annotation consistency and improve ML dataset quality tools such as Labelbox and CVAT to perform image and text annotation for AI models

PROJECTS

Rock Classification (Research Project) Link

* Analyzed 10,000 rock data and implemented classification models, Support Vector Machine, Random Forest, Gradient boost and Adaboost to predict the classification of rocks leveraging pandas, sklearn, matplotlib and seaborn

* Support Vector Machine gave the best model with a Test Set Accuracy score over 86%

Web Application for Aphasic People (*Senior Year Project*) Live Demo

* Built an assistive application using for aphasic people to memorize, pronounce words and facilitate improved phonetics

* Employed ReactJS to optimize application by increasing loading speed by 30%

Car Streaming Data Pipeline Link AWS Glue, Step Functions, Redshift, Aurora MySQL, S3, SQL

* Implemented robust data validation checks, improving pipeline reliability by 90%

* Designed efficient Redshift Upsert strategy using staging tables, reducing duplication errors by 80%

Music Streaming Pipeline Link Apache Airflow, Python, Amazon Redshift, S3, RDS, SQL

* Used Step Functions to orchestrate Glue jobs, ensuring error handling and stepwise execution

* Designed Computed key business metrics: Occupancy Rate, Average Listing Price, Repeat Customer Rate, and Top Performing Listings

POSITIONS OF RESPONSIBILITY

◦ Founded CODEHOY; A tech club in University of Mines and Technology (Former University) where I coordinated meetings and workshops to mentor, guide, and assist 500+ students in various specializations in tech.

CONFERENCES ATTENDED

◦ VI International Turkic World Congress on Science and Engineering, Azerbaijan-Baku, 2024

CERTIFICATION

◦ Diploma in Using Python For Data Science (Alison)

◦ Unsupervised & Supervised Machine Learning (Regression & Classification) (Coursera)