Michel Boris Emel Manfouo

+237 659 398 858 | michel.emel@aims-cameroon.org | LinkedIn | GitHub | Portfolio

Education

Junior Data Engineer, Data Scientist, and Machine Learning Engineer with a strong mathematical foundation. Proficient in Python, SQL, R, JavaScript, and tools like Apache Hadoop, Spark, Airflow, and Databricks. Skilled in building scalable data pipelines on GCP, AWS, and Azure, and working with Docker and Kubernetes. Expertise in TensorFlow, Scikit-learn, PyTorch, XGBoost, and deploying LLMs. Experienced in data visualization with Tableau, Power BI, and Matplotlib, and managing databases like PostgreSQL, MySQL, and MongoDB. Continuously learning and committed to delivering high-quality results.

Skills

Languages: Java, Python, AppScript, JavaScript, TypeScript, SQL, NoSQL MATLAB, R

Technologies: MySQL, PostgreSQL, Cassandra, Snowflake, Spark, Airflow, Databricks, BigQuery, MongoDB, Git, Docker, AWS, Kubernetes, GCP, Kafka, RabbitMQ, OpenCV, PyTorch, TensorFlow

Methodologies: OOP, Functional Programming, DevOps, CI/CD

Projects

GDP Data ETL Pipeline | *GitHub*

- Developed an efficient ETL pipeline to extract GDP data from Wikipedia, transform it into a consistent format, and load it into CSV and SQLite for comprehensive analysis.
- Implemented robust data handling procedures to ensure data accuracy and accessibility, facilitating seamless integration and analysis.

Largest Banks Data Extraction | *GitHub*

- Built an ETL pipeline using Python, BeautifulSoup, Pandas, and SQLite to extract, transform, and load global bank data with multi-currency market capitalization.
- Conducted data analysis with SQL queries and implemented logging to monitor ETL processes.

Obesity Prediction using Body Mass Index Dataset | *GitHub*

- Using preprocessing and advanced supervised learning techniques to predict obesity, emphasizing model interpretability, optimization.
- Developed expertise in feature engineering and model optimization to strengthen threat detection, anomaly identification, and incident response.

COVID-19 Prediction Model | *GitHub*

- Using preprocessing and advanced supervised learning techniques to predict obesity, emphasizing model interpretability, optimization.
- Developed expertise in feature engineering and model optimization to strengthen threat detection, anomaly identification, and incident response.

Experience

City Information Management System — Limbe, Cameroon

Data Engineer — *July 2024 – Present, Internship*

- Played a key role in the design and implementation of the Limbe City Information Management System (LCIMS) for property-related data management.
- Assisted in developing software modules for data collection, house numbering, and tax identification using Python, Java, and JavaScript.
- Leveraged MySQL and PostgreSQL for database management, ensuring efficient data storage and retrieval, collaborated on designing a revenue management dashboard
- Worked with web development frameworks such as Django, Flask, and React to build and test application components, and integrated APIs for third-party services.
- Provided technical support during system deployment and conducted training sessions for end-users, engaged in data analysis and report generation to enhance decision-making processes.

Certificates

Data Engineer Essentials *August 23, 2024*
Credential link

Python Project for Data Engineering *August 9, 2024*
Credential link

Databricks for Data Engineering *August 25, 2024*

Spark, Hadoop, and Snowflake for Data Engineering *August 28, 2024*
Credential link

Generative AI for Everyone *November 21, 2023*
Credential link

References

References available upon request.