PHAM QUOC HUY
Data Engineer

📞 +84 943884436 ✉ phamhuy1032003@gmail.com 📍 Tan Binh District, Ho Chi Minh city

🌐 https://www.linkedin.com/in/huy-pham-2a598026a/ 🔗 https://github.com/HuyPham103code/

Objects

Data Engineer with expertise in Python, SQL, and data pipeline development, seeking to apply technical skills to build efficient data infrastructure and support data-driven decision-making processes.

Skills

Technical skills
- Python, Golang, SQL
- PostgreSQL, Snowflake, Cassandra
- Spark, Kafka, Airflow, DBT
- AWS, Google Cloud

Tool skills
- Docker
- Power BI
- GitHub
- Visual Studio

Soft skills
- Problem-Solving
- Communication
- Fluent in English

Experience

Data Engineer Intern
Combros Company

Ho Chi Minh City | June 2024 – October 2024

- Analyzed data for labeling plant growth stages using machine learning models.
- Created a data pipeline to automate data flow from real-time sensors, storing data in Kafka and processing it using Spark, before saving it in Cassandra for further analysis.
- Utilized this processed data to train predictive models (K-means, GCN), and then deployed the models in a web application, integrating .pkl files on a server for client use.
- Generated reports on insights and model performance for stakeholders.
- Technologies: Docker | Python | Power BI | Machine Learning | Spark | Kafka | Cassandra | GCN

Personal Projects

● Snowflake Data Pipeline Automation   05/2024

This project automates the process of loading CSV data into a Snowflake database using Apache Airflow and dbt. By utilizing Snowflake's sample datasets, it streamlines data ingestion, transformation, and management in a cloud-based environment, enabling seamless and efficient data workflows.

Source: https://github.com/HuyPham103code/data_pipeline_airflow_snowflake

- Technologies: Apache Airflow | dbt | Snowflake | Python | Docker

● Data Pipeline for Automated Machine Learning Tasks   03/2024

Developed an end-to-end data pipeline for automating machine learning tasks, covering data extraction, transformation, and model training.

Source: https://github.com/HuyPham103code/data_pipe_airflow

- Technologies: Apache Airflow | Python | Docker | Pandas | matplotlib | scikit-learn

Education

● HCM Open University (HCMCOU) *2021 - present*
- Computer Science
- GPA: 3.1/4

● SQL (Intermediate) Certificate *2024*
- certificates