Tran Quoc Trung

Data Engineer

Portfolio: tranquoctrung.id.vn

tranquoctrung270302@gmail.com • 0917 802 192 • /in/tranquoctrung2703 • github.com/TrungTechh

Education

VNU-HCM University of Science   **Oct 2020 - Aug 2024**

Bachelor's degree Computer Science

GPA: 8.3/10

Skills

Programming	Python**,** SQL, C++
Tools	Apache Hadoop, PySpark, Airflow, Kafka, Power BI, Apache Superset, Docker, Git
DBMS	ClickHouse, Postgres, Microsoft SQL Server
Cloud	Microsoft Azure (ADLS, Azure Factory, Azure Functions, Azure Synapse, Azure SQL Database)
Languages	Proficient in English reading and communication in a work environment

Work Experience

7-Eleven - Fresher Data Engineer | Full-time   **Sep 2024 - Present**

- Collaborate with cross-functional teams to address and fulfill their data requirements.
- Develop and maintain real-time streaming pipelines that consume data from Kafka and transform it by ClickHouse.
- Design and implement Airflow jobs to process large datasets in batch, including automated notifications to Google Chat for job errors, missing data, or inactive Kafka consumer groups.
- Utilize App Script and n8n to integrate with external services such as Google, Jira for seamless data flow and automation.
- Developed a chatbot for generating insights from Superset data using OpenAI API.
- Enabled real-time replication from PostgreSQL using logical replication and Airbyte CDC.
- Technologies: **ClickHouse, Airflow, Kafka, App Script, n8n, Apache Superset, OpenAI API, Airbyte**

TC Data - Trainee Data Engineer | Full-time   **Jun 2024 - Sep 2024**

- Develop modules using SQL Server to help stakeholders meet their data needs.
- Engage in discussions with stakeholders to thoroughly verify and understand the issues present in the data.
- Designed automate ingest file from outlook and store into sharepoint using Power Automate.
- Technologies: **SQL Server, T-SQL, Power Automate**

Inter-K JSC - Data Engineer Intern | Full-time   **Nov 2023 - Jan 2024**

- Converted approximately 100 SAP crystal reports into paginated reports using PowerBI Report Builder.
- Technologies: **PowerBI Report Builder**, **Crystal Reports, DAX, Gitlab**

Technical Projects

Movies ETL Pipeline (github.com/TrungTechh/IMDBMovieETL)   **May 2024 - Jun 2024**

This project aims to utilize Azure to create a film scraping pipeline from the IMDB website for data preparation for analysis.

- Created an auto-scaling crawler using Azure Functions, deployed on Azure App Service.
- Established a Spark streaming job in Azure Databricks for proficient and automated data processing.
- Provided insightful visualizations and actionable insights on movie industry trends utilizing PowerBI.
- Technologies: **Azure Functions, Azure Databricks, Azure DLS, Spark Streaming, PowerBI**
- Demo video: https://www.youtube.com/watch?v=yB6MsZQX9TE

Streaming Data Pipeline (github.com/TrungTechh/Weather-and-Daily-Life)   **Jan 2024 - Apr 2024**

Developing an application that can perform real-time analysis of the weather conditions, service quality, and traffic status in Ho Chi Minh City.

- Used YOLOv8 to detect 5+ vehicle types in traffic footage from Ho Chi Minh City's official site.
- Engineered solution to efficiently process real-time traffic data using Kafka, which increased data ingestion speed by 20%.
- Implemented data visualization dashboards in Power BI to present actionable insights to stakeholders.
- Technologies: **Python, Kafka, Power BI, PySpark, Cassandra, Airflow, Flask, Docker**

Certifications

TOEIC (Listening and Reading)**: 815**

DP-900: **Azure Data Fundamentals**