



KỸ SƯ DỮ LIỆU

NGÔ TRUNG QUANG

MỤC TIÊU NGHỀ NGHIỆP

Tôi muốn trở thành chuyên gia về tối ưu hóa truy vấn SQL, sử dụng chỉ mục thông minh và partitioning trong các hệ thống dữ liệu lớn nhằm rút ngắn thời gian phản hồi truy vấn.

THÔNG TIN CÁ NHÂN

07/06/1984

Hà Nội

thaolinh252512@gmail.com

0762133680

www.website.com

HỌC VẤN

- Phân tích và xử lý dữ liệu lớn tại Đại học Phenikaa

KỸ NĂNG

- PostgreSQL

- Airflow

- Snowflake

KINH NGHIỆM LÀM VIỆC

- **JUNIOR DATA ENGINEER** TẠI RETAILTECH (2020-2021)

+ PHỐI HỢP XÂY DỰNG HỆ THỐNG PHÂN TÍCH TỒN KHO THEO THỜI GIAN THỰC

+ XỬ LÝ DỮ LIỆU STREAMING TỪ KAFKA VÀ LƯU TRỮ VÀO BIGQUERY

+ XÂY DỰNG CÁC VIEW LOGIC TRÊN DBT PHỤC VỤ BI DASHBOARD

- **BIG DATA ENGINEER** TẠI LOGIAI (2020-2022)

+ XÂY DỰNG PIPELINE THU THẬP VÀ XỬ LÝ DỮ LIỆU

SỞ THÍCH

- Sưu tầm sách lập trình

NGƯỜI GIỚI THIỆU

- Bà Trần Thị Mai (Technical Manager - Công ty CloudData) - mai.tran@cloudata.vn - 0933666777

VẬN CHUYỂN TỪ THIẾT BỊ IOT

+ SỬ DỤNG SPARK STREAMING ĐỂ XỬ LÝ DỮ LIỆU REAL-TIME

+ KẾT NỐI HỆ THỐNG VÀO ELASTICSEARCH VÀ PHÁT TRIỂN DASHBOARD GIÁM SÁT

- **CLOUD DATA ENGINEER** TẠI SMARTBANK (2022-2023)

+ THIẾT KẾ KIẾN TRÚC DỮ LIỆU TRÊN AWS VỚI S3, GLUE VÀ REDSHIFT

+ TỰ ĐỘNG HÓA VIỆC INGESTION LOG GIAO DỊCH NGÂN HÀNG TỪ HỆ THỐNG CORE

+ THAM GIA ĐÁNH GIÁ BẢO MẬT VÀ CHUẨN HÓA SCHEMA DỮ LIỆU TÀI CHÍNH

- **DATA ENGINEER** TẠI CÔNG TY DATAFLOW VIỆT NAM (2021-2023)

+ THIẾT KẾ VÀ TRIỂN KHAI PIPELINE ETL SỬ DỤNG APACHE AIRFLOW

+ TÍCH HỢP DỮ LIỆU TỪ NHIỀU NGUỒN NHƯ MYSQL, GOOGLE ANALYTICS VÀ API BÊN THỨ BA

+ TỐI ƯU HOÁ QUY TRÌNH XỬ LÝ BẰNG SPARK TRÊN GCP DATAPROC

DANH HIỆU VÀ GIẢI THƯỞNG

- **2022** - Top 3 nhân viên cải tiến hệ thống ETL tại SmartRetail

CHỨNG CHỈ

- **2020** - AWS Certified Data Analytics – Specialty

- **2023** - Designing Data-Intensive Applications – O'Reilly Certification Program

- **2020** - Google Cloud Professional Data Engineer Certificate

- **2023** - Modern Data Engineering with dbt – dbt Labs

HOẠT ĐỘNG

- Thành viên nhóm tư vấn kiến trúc dữ liệu tại Công ty SmartData (2021 - 2022)

+ Đề xuất mô hình lưu trữ dữ liệu dạng columnar cho truy vấn phân tích nhanh.

+ Tham gia thiết kế cấu trúc bảng dữ liệu trong hệ thống warehouse.

+ Xây dựng best practice về đặt tên và phân vùng dữ liệu.

DỰ ÁN

- Realtime Analytics cho hệ thống bán lẻ toàn quốc (Big Data Engineer, Retail360) 2023

Xây dựng hệ thống xử lý dữ liệu bán hàng theo thời gian thực để hỗ trợ ra quyết định tức thì.

- + Sử dụng Kafka để thu thập dữ liệu từ các chi nhánh toàn quốc
- + Xử lý dữ liệu streaming bằng Spark Structured Streaming
- + Gửi dữ liệu về Redshift và hiển thị trên dashboard BI

- ETL tự động hóa cho nền tảng tài chính (ETL Developer, FinTrack) 2021

Tự động hoá quá trình tổng hợp báo cáo từ nhiều hệ thống tài chính nội bộ.

- + Viết ETL pipeline bằng Python và Pandas
- + Tích hợp với hệ thống email để gửi báo cáo tự động hàng tuần
- + Quản lý lịch ETL với crontab và giám sát bằng Prometheus

