



KỸ SƯ DỮ LIỆU

HOÀNG NGÀ DƯƠNG

MỤC TIÊU NGHỀ NGHIỆP

Tôi hướng tới mục tiêu trong 3 năm tới trở thành một Senior Data Engineer, có khả năng thiết kế toàn bộ kiến trúc dữ liệu, dẫn dắt đội nhóm và đề xuất cải tiến cho chiến lược dữ liệu doanh nghiệp.

THÔNG TIN CÁ NHÂN

20/11/1994

Hà Nội

thaolinh252512@gmail.com

0921332167

www.website.com

KINH NGHIỆM LÀM VIỆC

- **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

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

HỌC VẤN

- Toán - Tin ứng dụng tại Đại học Kinh tế Quốc dân

KỸ NĂNG

- Apache Spark

- Python

SỞ THÍCH

- Tham gia hội thảo công nghệ
- Du lịch
- Đi bộ đường dài

+ 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

NGƯỜI GIỚI THIỆU

- Bà Nguyễn Thị Lệ (Analytics Lead - Công ty InsightTech) -
le.nguyen@insighttech.vn - 0966999888

- Ông Lê Hoàng Dương (Head of Engineering - Công ty FinData) -
duong.le@findata.vn - 0988999123

- Bà Phạm Thanh Hương (Data Operations Manager - Công ty ReportPro) -
huong.pham@reportpro.vn - 0912111222

- Bà Đỗ Ngọc Trang (Senior Data Engineer - Công ty SmartRetail) -
trang.do@smartretail.vn - 0922333444

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

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

+ XÂY DỰNG PIPELINE THU THẬP VÀ XỬ LÝ DỮ LIỆU 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

- **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

DANH HIỆU VÀ GIẢI THƯỞNG

- **2023** - Giải thưởng 'Kiến trúc dữ liệu sáng tạo' do CTO trao tặng tại DataWare
- **2022** - Top 3 nhân viên cải tiến hệ thống ETL tại SmartRetail
- **2022** - Vinh danh 'Nhân sự triển khai hệ thống streaming real-time đầu tiên' tại LogiTech
- **2021** - Kỹ sư dữ liệu xuất sắc quý I tại Công ty DataBridge

CHỨNG CHỈ

- **2023** - Designing Data-Intensive Applications – O'Reilly Certification Program
- **2021** - SQL for Data Engineering – Datacamp
- **2022** - Big Data Engineering with Apache Spark – edX (BerkeleyX)
- **2021** - Streaming Data with Apache Kafka – Confluent Academy

HOẠT ĐỘNG

- Thành viên nhóm kỹ thuật dữ liệu tại Dự án E-Government (2023)

+ Thiết kế hệ thống thu thập dữ liệu hành chính từ nhiều bộ ngành.

- + Thiết lập hệ thống kiểm tra chất lượng dữ liệu tự động.

- + Triển khai pipeline đồng bộ dữ liệu hằng ngày với độ trễ thấp.

- Thành viên nhóm phát triển hệ thống phân tích log tại Công ty SecureLog (2021)

- + Xây dựng pipeline thu thập và phân tích log từ nhiều nguồn hệ thống.

- + Triển khai hệ thống cảnh báo bất thường dựa trên mô hình thống kê.

- + Sử dụng ELK stack để trực quan hóa dữ liệu log.

DỰ ÁN

- Data Lake cho hệ thống quản trị khách hàng (CRM) (Data Engineer, CRMPro) 2022

Xây dựng nền tảng lưu trữ dữ liệu tập trung phục vụ phân tích hành vi khách hàng.

- + Tạo pipeline ingestion từ Salesforce, Google Ads và Facebook API

- + Lưu trữ dữ liệu theo mô hình phân vùng S3 Data Lake

- + Sử dụng Airflow để lập lịch và monitor luồng dữ liệu hàng ngày

