



KỸ SƯ DỮ LIỆU

TRẦN PHƯƠNG TÚ

MỤC TIÊU NGHỀ NGHIỆP

Tôi định hướng trở thành một Kỹ sư Dữ liệu chuyên nghiệp, có khả năng xây dựng các pipeline xử lý dữ liệu mạnh mẽ và tối ưu nhằm phục vụ các hệ thống phân tích và học máy trong doanh nghiệp.

THÔNG TIN CÁ NHÂN

24/09/1989

Hà Nội

thaolinh252512@gmail.com

0724512710

www.website.com

HỌC VẤN

- Cơ sở dữ liệu và khai phá dữ liệu tại Đại học Khoa học Tự nhiên – ĐHQG Hà Nội
- Phân tích và xử lý dữ liệu lớn tại Đại học Phenikaa

KỸ NĂNG

- NoSQL (MongoDB, Cassandra)
- Docker
- AWS (S3, EMR, Glue)

KINH NGHIỆM LÀM VIỆC

- **ETL DEVELOPER** TẠI FINDATA (2019-2021)

+ PHÁT TRIỂN VÀ BẢO TRÌ LUỒNG DỮ LIỆU ETL SỬ DỤNG TALEND VÀ PYTHON

+ VIẾT SCRIPT KIỂM TRA CHẤT LƯỢNG DỮ LIỆU VÀ CẢNH BÁO KHI DỮ LIỆU SAI LỆCH

+ TÍCH HỢP HỆ THỐNG PHÂN QUYỀN DỮ LIỆU THEO VAI TRÒ NGƯỜI DÙNG

DANH HIỆU VÀ GIẢI THƯỞNG

- **2023** - Giải thưởng 'Kiến trúc dữ liệu sáng tạo' do CTO trao tặng tại DataWare

SỞ THÍCH

- Chụp ảnh
- Thể thao
- Tham gia cộng đồng lập trình
- Học ngoại ngữ
- Viết blog kỹ thuật

NGƯỜI GIỚI THIỆU

- Bà Đỗ Ngọc Trang (Senior Data Engineer - Công ty SmartRetail) - trang.do@smartretail.vn - 0922333444
- Bà Lê Thùy Linh (Data Infrastructure Manager - Công ty CloudOps) - linh.le@cloudops.vn - 0944555666

CHỨNG CHỈ

- **2022** - Data Engineering on Google Cloud Platform - Coursera
- **2020** - AWS Certified Data Analytics - Specialty
- **2020** - Google Cloud Professional Data Engineer Certificate

HOẠT ĐỘNG

- Tình nguyện viên hỗ trợ tổ chức Data Day tại Vietnam Data Community (2022)

- + Thiết lập demo pipeline dữ liệu thời gian thực bằng Kafka.
- + Hỗ trợ phần trình diễn luồng dữ liệu streaming giữa các node Spark.
- + Giới thiệu cấu trúc cơ bản của hệ thống Data Lake.

- Cộng tác viên dự án dữ liệu mở tại OpenData Vietnam (2020)

- + Thu thập và chuẩn hóa dữ liệu dân số từ các tỉnh thành.
- + Tích hợp dữ liệu vào hệ thống cơ sở dữ liệu MongoDB phục vụ phân tích.
- + Tạo dashboard đơn giản để trực quan hóa dữ liệu bằng Plotly Dash.

- 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.

- Tình nguyện viên hỗ trợ nghiên cứu học thuật tại Khoa Công nghệ Thông tin (2020)

- + Chuẩn hóa dữ liệu nghiên cứu sinh viên từ nhiều năm học.
- + Phân tích và tổng hợp xu hướng đăng ký môn học.
- + Viết báo cáo kỹ thuật trình bày cho giảng viên hướng dẫn.

DỰ ÁN

- 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

