|  |  |  |
| --- | --- | --- |
| Decorative |  | Kỹ sư Dữ liệu |
|  | Đặng Tú LanMỤC TIÊU NGHỀ NGHIỆP **Tôi định hướng phát triển các kỹ năng DevOps cho dữ liệu như CI/CD pipeline cho hệ thống ETL, quản lý version dữ liệu bằng DVC, và monitor hệ thống bằng Prometheus/Grafana.** |
| THÔNG TIN CÁ NHÂN  21/07/1997  Hà Nội  thaolinh252512@gmail.com  0860115879  www.website.com  HỌC VẤN - Cơ sở dữ liệu và khai phá dữ liệu tại Đại học Khoa học Tự nhiên – ĐHQG Hà Nội KỸ NĂNG  - Python  - Monitoring với Prometheus/Grafana SỞ THÍCH - Chụp ảnh  - Đi bộ đường dài  - Chơi đàn guitar  - Tham gia hackathon  - Tập gym NGƯỜI GIỚI THIỆU - Bà Đỗ Ngọc Trang (Senior Data Engineer - Công ty SmartRetail) - trang.do@smartretail.vn - 0922333444  - Bà Nguyễn Thị Lệ (Analytics Lead - Công ty InsightTech) - le.nguyen@insighttech.vn - 0966999888 |  | 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- **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 sátDANH HIỆU VÀ GIẢI THƯỞNG - **2020** - Giải thưởng 'Pipeline ổn định nhất năm' tại TechFlow  - **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 - **2020** - Google Cloud Professional Data Engineer Certificate  - **2021** - Streaming Data with Apache Kafka – Confluent Academy  - **2021** - IBM Data Engineering Professional Certificate – Coursera  - **2023** - Designing Data-Intensive Applications – O’Reilly Certification Program HOẠT ĐỘNG **- 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.  **- 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 **- 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  **- Kiến trúc dữ liệu cho nền tảng giáo dục trực tuyến (Cloud Data Engineer, LearnHub) 2022**  Chuẩn hoá kiến trúc lưu trữ và xử lý dữ liệu học viên để phục vụ dashboard học tập.  + Thiết kế hệ thống lưu trữ với BigQuery và Data Studio  + Xây dựng Dataflow jobs để xử lý dữ liệu sự kiện học tập  + Tạo luồng dữ liệu từ Firebase tới Google Cloud  **- 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 |
|  |  |  |