|  |  |  |
| --- | --- | --- |
| Decorative |  | Kỹ sư Dữ liệu |
|  | Bùi Thị HàMỤC TIÊU NGHỀ NGHIỆP **Tôi mong muốn phát triển sự nghiệp với vai trò Data Engineer tại môi trường sử dụng dữ liệu lớn, nơi tôi có thể vận dụng Apache Spark, Kafka và Airflow để xử lý và vận hành pipeline dữ liệu theo thời gian thực.** |
| THÔNG TIN CÁ NHÂN  08/08/1995  Hà Nội  thaolinh252512@gmail.com  0766967583  www.website.com  HỌC VẤN - Khoa học dữ liệu tại Đại học Công nghệ Thông tin – ĐHQG TP.HCM KỸ NĂNG  - Data Lake Architecture  - Snowflake  - Monitoring với Prometheus/Grafana  - AWS (S3, EMR, Glue) SỞ THÍCH - Chụp ảnh  - Thử nghiệm công nghệ mới  - Xem phim khoa học viễn tưởng  - Học ngoại ngữ NGƯỜI GIỚI THIỆU - Ông Nguyễn Quốc Huy (Data Platform Lead - Công ty DataBridge) - huy.nguyen@databridge.vn - 0905123456 |  | 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- **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- **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 chínhDANH HIỆU VÀ GIẢI THƯỞNG - **2021** - Kỹ sư dữ liệu xuất sắc quý I tại Công ty DataBridge CHỨNG CHỈ - **2021** - SQL for Data Engineering – Datacamp  - **2022** - ETL and Data Pipelines with Shell, Airflow and Kafka – Coursera 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.  **- 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 **- 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 |
|  |  |  |