|  |  |  |
| --- | --- | --- |
| Decorative |  | Kỹ sư Dữ liệu |
|  | Phạm Nam QuangMỤC TIÊU NGHỀ NGHIỆP **Tôi định hướng tạo ra nền tảng dữ liệu mở (open data platform) cho các bộ phận nội bộ dễ dàng truy cập, khai thác và sử dụng dữ liệu theo cơ chế phân quyền rõ ràng.** |
| THÔNG TIN CÁ NHÂN  24/06/1990  Hà Nội  thaolinh252512@gmail.com  0906027696  www.website.com  HỌC VẤN - Kỹ thuật phần mềm tại Học viện Công nghệ Bưu chính Viễn thông KỸ NĂNG  - NoSQL (MongoDB, Cassandra) SỞ THÍCH - Du lịch  - Tham gia hội thảo công nghệ  - Thể thao 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  - Bà Trần Thị Mai (Technical Manager - Công ty CloudData) - mai.tran@cloudata.vn - 0933666777  - Ông Nguyễn Quốc Huy (Data Platform Lead - Công ty DataBridge) - huy.nguyen@databridge.vn - 0905123456  - Ông Lê Hoàng Dương (Head of Engineering - Công ty FinData) - duong.le@findata.vn - 0988999123 |  | 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 DataprocDANH HIỆU VÀ GIẢI THƯỞNG - **2021** - Giải thưởng 'Dự án phân tích dữ liệu xuất sắc' tại FinData  - **2020** - Giải thưởng 'Pipeline ổn định nhất năm' tại TechFlow  - **2022** - Top 3 nhân viên cải tiến hệ thống ETL tại SmartRetail  - **2020** - Top 5 kỹ sư xử lý dữ liệu lớn trong cuộc thi nội bộ AI Challenge CHỨNG CHỈ - **2020** - AWS Certified Data Analytics – Specialty  - **2021** - IBM Data Engineering Professional Certificate – Coursera  - **2023** - Modern Data Engineering with dbt – dbt Labs 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. 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  **- Data pipeline xử lý log giao dịch hệ thống ngân hàng (Data Pipeline Engineer, BankDataTech) 2020**  Lưu trữ, phân tích và phát hiện bất thường từ log giao dịch ATM và mobile banking.  + Sử dụng Fluentd thu thập log và gửi về Kafka  + Phân tích real-time qua Spark Streaming  + Lưu dữ liệu đã xử lý vào ElasticSearch và trực quan hoá bằng Kibana  **- 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 |
|  |  |  |