DUY NGUYEN

LEAD DATA ENGINEER

☎ +84 399360638

✉ duynguyenngoc@hotmail.com

⌖ Ho Chi Minh city

🌐 duynguyenngoc.com

PROFILE

As a seasoned Senior Data Engineer with a robust background in Software Engineering, I possess a wealth of expertise in Automation, Backend Development, Artificial Intelligence, and Data Processing & Visualization. I am passionate about leveraging my skills to solve complex problems and drive innovation in the field of data.

SKILLS

- Python, SQL, ReactJS
- Data Warehouse, Lake
- Extract, Transform, Load (ETL)
- Document Processing
- Airflow, Spark, Kafka, Superset
- CI/CD
- Kubernetes, Docker
- Problem-Solving
- Strong Communication

EDUCATION

BACHELOR'S DEGREE OF ELECTRICAL (TELECOMMUNICATION)

Ho Chi Minh city University of technology

Programme de Formation d'Ingénieurs d'Excellence au Vietnam (2019)

EXPERIENCE

SENIOR DATA ENGINEER

VieON (September 2023 - Present)

- Lead migration of current Data Platform (HDFS, Spark, Superset, etc.) to Kubernetes. Ensure minimal downtime and optimal performance with regular maintenance. Establish monitoring and alert systems.
- Lead 4-person Data Engineer team, providing guidance and fostering innovation. Align team activities with project objectives.
- Architect's scalable platform (Trino, MinIO, Kafka, Spark, Spark Streaming, Kafka, etc.) for 5+ companies, handling 2+ TB data per day. Collaborate with stakeholders for tailored solutions.
- Maintain and enhance features for Recommendation API. Collaborate with data scientists for scalable solutions.

SENIOR DATA ENGINEER

DIGI-TEXX (March 2022 - September 2023)

- Design and build a scalable data pipeline that can handle 1TB images per day, including a data warehouse and data lake based on S3, Hbase, Hadoop, Hive, Superset, and Airflow.
- Design and Implement AI Product on Big Data Infastructure, including but not limited to: ML, DL, HDFS, Spark, Kafka.
- Implement microservices in AI systems with scalable, efficient code to reduce 70% operational costs.
- Implement Image Classification, Image Denoising, Field Detection and Optical Character Recognition models using AI techniques.
- Managing AI engineering research team.

DATA ENGINEER

DIGI-TEXX (September 2020 - March 2022)

- Developed an ETL/ELT system collect more than 100 data sources to replace an existing data pipeline.
- Collaborated with data scientists and product engineers to develop and deploy solutions to customer problems.
- Analytic Data Modeling (Star Schema, SCD's, and Lookups)
- Increased the efficiency of the data fetching by approximately 30% using query optimization, indexing and partition.

BACKEND DEVELOPER / AI ENGINEER

Amaris Consulting (December 2019 - September 2020)

- Proposing solutions and technical design of social ios backend applications.
- Deploy REST-FULL API endpoints with FastAPI on AWS services.
- Build a product recommendation system based on product description with Deep Learning.
- Build bots to automatically collect data from mutiple source.
- Performing analysis to assess quality and meaning of data.

AUTOMATION TESTER

TMA Solution (April 2019 - October 2019)

- Deploy automation framework with selenium on cross-platform android, ios, window, macos and web.
- Research and propose solutions to optimize automation framework.
- Executing telecom system test cases on multiple platforms through automation framework.
- detect bugs, report them to developers and help them reproduce the bug.

MACHINE LEARNING ENGINEER JUNIOR

Data Science Lab (August 2018 - February 2019)

- Provide technical solutions to detect accident victims in inaccessible terrains such as cracks, large fires, etc. through image analysis
- Development of victim detection model with Deep Learning based on Tensorflow and Python
- Preprocessing data and training model.
- Build a real-time victim detection system with Google Firebase.