

HOANG VAN TRUONG

DATA ENGINEER

0392582089 | truonghoang150602@gmail.com | [linkedin](#) | [portfolio](#)

SUMMARY

Passionate and disciplined Data Engineer eager to learn and grow in a professional environment. Skilled in databases, data warehouses, data lakes, and data processing. Ready to contribute with enthusiasm and a quick learning ability, aiming to build and optimize efficient data systems.

EDUCATION

Hanoi University of Science and Technology

Vietnam-Japanese Information Technology (AI and Big Data)

Oct 2020 - Sep 2024

Bachelor's degree (CPA: 3.23/4)

EXPERIENCE

FPT Software - Data Engineer

Oct 2024 - Present

- Designed and implemented ETL data pipelines for a chatbot system to extract and process information from a large volume of unstructured files (.pdf, .docx, .pptx) containing complex content such as graphs, charts, and images.
- Development for the DOCX and PDF data pipeline, ensuring accurate parsing and structured output, while providing support for other file formats.
- Collaborated closely with the AI team to refine and restructure the vector database, significantly improving the chatbot's response accuracy and relevance.

Nissan Automotive Vietnam - Data Engineer (on-site)

Oct 2024 - Apr 2025

- Optimized a high-load MySQL database (10M+ records) by redesigning the structure, enhancing complex SQL query performance, and improving read/write operations for better efficiency and scalability.
- Developed python tool to extract and parse data from technical files, enabling efficient analysis and visualization for stakeholders.

Tiki - Research Collaborator (part-time)

Dec 2023 - Mar 2024

- Processed and analyzed a large volume of customer reviews from Tiki products using Apache Spark for distributed data preprocessing and cleaning.
- Built data preparation pipelines with Spark MLlib's Transformers and Encoders to train machine learning models for sentiment and relevance analysis.
- Visualized key insights through data visualization tools to support business decision-making and drive product improvement strategies.

Koolsoft E-learning - Full-stack developer (part-time)

Oct 2023 - Jan 2024

- Developed a dynamic website that allows users to practice and complete a variety of exercises across different formats.
- Built the platform using the MERN stack (MongoDB, Express.js, React.js, Node.js) to ensure scalability, responsiveness, and a smooth user experience.

TECHSTACK

- Programing Languages:** Python, Scala, JavaScript
- Data Processing Frameworks:** Pandas, Apache Spark, Apache Airflow, Apache Kafka, etc
- Data Warehousing & Databases:** MySQL, PostgreSQL, DynamoDB, S3, etc
- Others:** Git, Docker

ADDITIONAL INFORMATION

- Languages:** English (Toeic 550), Japanese (JLPT N3)
- Certifications:** AWS Data Engineer Associate (DEA - C01)