

Truong Tien Anh

Data Engineer

+ (84) 373 097 346 | ✉ truongtienanh16@gmail.com | [in LinkedIn](#) | [GitHub](#) | [My Website](#)

EDUCATION

University of Science - VNUHCM

Bachelor of Science in Data Science and Computer Science

Oct. 2022 – Present

Current GPA: 3.7/4.0

Language Skills

(Conversational, able to work in English-speaking environment)

TOEIC: 810/990, **VSTEP:** B2

CAREER OBJECTIVE

Final-year Computer Science student at VNUHCM - University of Science with strong interest in data engineering, large-scale processing, and building reliable pipelines. Seeking an internship or entry-level role as a **Data Engineer** to apply my skills in designing and optimizing workflows while growing expertise in big data technologies.

SKILLS

Languages: Python, SQL, JavaScript, C/C++

Databases: SQL Server, PostgreSQL, MongoDB, Qdrant

Full-Stack Development: Express.js, Node.js, Next.js, HTML, CSS, Pug, FastAPI

Machine Learning: RAG, Chatbot, LLM, LangChain, TensorFlow, Scikit-learn

Big Data: Apache Hadoop, Spark, Airflow, Kafka, Docker

Cloud: GCP, AWS (S3, Glue, Redshift, Athena), Git/GitHub, MongoDB Atlas

EXPERIENCE

Nexlab Technology – AI Engineer Intern

July 2025 – Oct 2025

- **1. Scan CV Project:** Developed an AI-driven system to scan, extract, and enhance CV data for streamlined recruitment. Integrated advanced retrieval techniques (RAG, Text2Sql, etc.) for efficient candidate matching.
- **2. Sales Agent Development:** Contributed to building a chatbot for querying retail data on GCS, enabling real-time insights and improving decision-making in retail management. Integrated MacSQL, a multi-agent system, for optimized SQL query generation and execution across distributed databases, enhancing data retrieval efficiency.

PROJECTS

End-to-End MovieDB | [GitHub Link](#)

2/2025 – 3/2025

- **Team size:** Solo project
- **Key Technologies:** Apache Airflow, Spark, Kafka, HDFS, PostgreSQL, Streamlit, Machine learning
- **Description:**
 - * Built an end-to-end pipeline to crawl, clean, and transform movie data into structured formats.
 - * Stored raw/processed data in HDFS and PostgreSQL; performed cleaning and analysis with Spark.
 - * Orchestrated ETL workflows with Kafka & Airflow for automation.
 - * Deployed a Streamlit app for movie price prediction and personalized recommendations.

Mobile AWS Pipeline Engineering | [GitHub Link](#)

3/2025 – 4/2025

- **Team size:** Solo project
- **Key Technologies:** AWS (S3, Glue, Athena, Redshift), Docker, Apache Kafka, Spark, PostgreSQL, Machine learning
- **Description:**
 - * Developed AWS-based pipeline for mobile recommendation and price prediction.
 - * Ingested streaming data via Kafka; processed in real time with Spark; automated cleaning with Glue.
 - * Queried insights using Athena/Redshift and visualized with dashboards.
 - * Optimized infrastructure for 30–50% faster data processing.

- **Team size:** 3
- **Key Technologies:** Kafka, Spark Structured Streaming, MongoDB, Docker, Airflow
- **Description:**
 - * Implemented a real-time pipeline to stream Bitcoin prices from Binance via Kafka.
 - * Computed moving averages and Z-scores with Spark Structured Streaming.
 - * Stored analytics in MongoDB for querying and visualization.
 - * Added anomaly detection logic to capture short-term price spikes.

ACHIEVEMENTS & AWARDS

Awarded scholarship from NinetyEight.	<i>May 2025</i>
HCMC AIChallenge 2024	<i>Oct 2024</i>
Ranked Top 30 in the AI Challenge 2024 organized by UIT.	<i>July 2024</i>
Awarded scholarship from Dong Hanh.	<i>Oct 2023</i>
Excellent Student Scholarship (HCMUS) , for the academic year 2022–2023.	<i>Mar 2023</i>
Awarded scholarship from VGSP, PwC, Vu A Dinh, ..etc.	<i>Apr 2024</i>

CERTIFICATIONS

VSTEP - B2 - Vang Lang University	<i>Aug 2025</i>
TOEIC - 810/990 - ETS Global	<i>July 2025</i>
Unsupervised Learning, Recommenders, Reinforcement Learning - Coursera, Deeplearning.ai	<i>Mar 2025</i>
Python for Data Science and Machine Learning Bootcamp - Udemy	<i>Mar 2025</i>
Supervised Machine Learning: Regression and Classification - Coursera, Deeplearning.ai	<i>Mar 2025</i>
The Data Science Course Bootcamp - Udemy	<i>Nov 2024</i>
SQL Advanced - Udemy	<i>Sep 2024</i>
Python for developer - Udemy	<i>Sep 2024</i>