



# Mai Van Tan

## Data Engineering Intern

+84 976 041 145  
tan.maivande@gmail.com  
linkedin.com/in/tanmaivan  
github.com/tanmaivan  
tanmaivan.vercel.app  
Ho Chi Minh City

## EDUCATION

University of Information Technology – VNUHCM  
Bachelor of Computer Science

Sep 2022 – Sep 2025  
GPA: 8.85/10

## SKILLS

- **Core Competencies:** Proficient in Python, SQL.
- **Data Engineering & Big Data:** Apache Spark, Apache Kafka, Debezium, dbt, Apache Airflow, Prefect.
- **Cloud, DevOps & Infrastructure:** Docker, Terraform, AWS, GCP, Linux, Git/GitHub.
- **Databases & Storage:** PostgreSQL, MySQL, DuckDB, MongoDB, Snowflake, BigQuery, MinIO.

## PROJECTS

### Data Science Salary Prediction

Nov 2024 – Feb 2025

- Collaborated on EDA and Modeling phases in a team of 4, analyzing 3,000+ job listings to identify key salary drivers.
- Benchmarked 6 regression algorithms and fine-tuned an XGBoost model to achieve a top  $R^2$  score of 0.83.
- **Tech Stack:** Python, Scikit-learn, Pandas.

### Hacker News Analytics Platform

Sep 2025 – Oct 2025

- Engineered a 100% Serverless ELT pipeline capable of continuous, unlimited ingestion from Hacker News API to BigQuery.
- Modeled data into a Star Schema to visualize Content Engagement, Author Influence, and Community Behavior via Looker Studio.
- Managed 10+ cloud resources purely via Terraform (IaC), ensuring a fully reproducible infrastructure environment.
- **Tech Stack:** Python, GCP: GCS/Big Query, Terraform, Prefect Cloud, dbt, Looker Studio

### Geo Weather Lake

Oct 2025 – Nov 2025

- Automated hourly ingestion for 34 administrative units in Vietnam, merging 30 days of historical backfill with real-time streams.
- Applied Geospatial Enrichment to visualize Pin Maps, Temperature/Rainfall trends, and regional comparisons on Metabase.
- Reduced analytical query time to milliseconds by leveraging DuckDB on top of Delta Lake storage.
- **Tech Stack:** Python, Docker, Airflow, dbt, MinIO, DuckDB, Metabase.

### SafeBank Data Platform

Nov 2025 – Dec 2025

- Architected a production-grade Medallion Architecture on Delta Lake, handling ACID transactions for 12+ banking tables.
- Achieved sub-minute latency via Debezium CDC and implemented SCD Type 2 for 100% compliance tracking.

- Designed Executive Dashboards to monitor Financial Performance, Risk Controls (NPL, Smurfing), and Customer Segmentation (RFM).
- **Tech Stack:** Python, Docker, Airflow, PostgreSQL, Apache Spark, Debezium, Apache Kafka, MinIO, Metabase

ACHIEVEMENTS

Academic Encouragement Scholarship

Awarded for Excellent Academic Performance

2022 – 2025

- Semester 2, Academic Year 2022–2023
- Semester 1, Academic Year 2023–2024
- Semester 1, Academic Year 2024–2025



Student of Five Merits

Recognized for excellence in Ethics, Academics, Fitness, Volunteering, and Integration

2022 – 2024

- Academic Year 2022–2023
- Academic Year 2023–2024

CERTIFICATIONS

- **TOEIC Listening & Reading:** 895/990 (ETS) 2025
- **TOEIC Speaking & Writing:** 330/400 (ETS) 2025
- **SQL (Advanced):** HackerRank Certified  2025
- **Problem Solving Using Computational Thinking:** Coursera  2024

SOCIAL ACTIVITIES

Study Board of Software Engineering

English Team Leader

2023 – 2024

- Led the English team in creating creative digital media content to promote learning.
- Organized internal training sessions and final exam review workshops for students.

Class Youth Union Committee

Secretary & Deputy Secretary

2023 – 2025

- Served as the primary liaison between the Faculty and the class, boosting student engagement in academic and social events.

UIT Job Fair Organizing Team

Event Support Member

2022 – 2024

- Assisted in organizing and managing logistics for the annual UIT Job Fair events.