

# Truong Tien Anh

+(84) 982 619 731 | [truongtienanh16@gmail.com](mailto:truongtienanh16@gmail.com) | [portfolio.com/trgtanh](https://portfolio.com/trgtanh) | [github.com/trgtanh04](https://github.com/trgtanh04)

## EDUCATION

---

**University of Science - VNUHCM**

*Bachelor of Science in Data Science and Computer Science*

Oct. 2022– Present

*Current GPA: 3.7/4.0*

## CAREER OBJECTIVE

---

I am a third year Computer Science student with a passion for data science, machine learning, and big data. I am looking for an internship or job opportunity as a Data Engineer or Data Scientist to apply my skills to real-world problems and grow through hands-on experience.

## SKILLS

---

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

**Data Science & ML:** Pandas, NumPy, Scikit-learn, TensorFlow, Matplotlib, Seaborn, Plotly

**Data Engineering:** Apache Hadoop, Spark, Airflow, Kafka, PostgreSQL

**Cloud & DevOps:** AWS (S3, Glue, Redshift, Athena), Docker, Git/GitHub, MongoDB Atlas, Neon Cloud

## CERTIFICATION & BLOG

---

**TOEIC:** 810/990

**My portfolio:** [portfolio.com/trgtanh](https://portfolio.com/trgtanh)

## PROJECTS

---

### End to End MovieDB

February 2025 – Present

- **Key Technologies:** Python, Apache Airflow, Spark, Kafka, HDFS, PostgreSQL, Streamlit
- **Description:**
  - \* Designed to support data processing and power a movie recommendation application.
  - \* Built a data pipeline to crawl and transform movie data into structured JSON.
  - \* Stored raw data in HDFS and used Spark for cleaning and basic analysis.
  - \* Kafka triggered ETL workflows; cleaned data was loaded into PostgreSQL.
  - \* Deployed PostgreSQL to Neon for cloud accessibility.
  - \* Implemented a model to predict movie prices and a Streamlit app for recommendations.
- **GitHub:** <https://github.com/trgtanh04/End-to-End-MovieDB-Data-Engineering>

### Mobile AWS Pipeline Engineering

March 2025 – Present

- **Key Technologies:** AWS (S3, Glue, Athena, Redshift), Docker, Apache Kafka, Apache Spark, Python, PostgreSQL
- **Description:**
  - \* Designed to power a phone recommendation system and predict mobile phone prices.
  - \* Boosted data processing speed by 30–50% using Kafka and AWS-based infrastructure.
  - \* Crawled mobile phone data, staged with Kafka, and logged events in PostgreSQL.
  - \* Used Spark for real-time stream processing and loaded outputs to S3.
  - \* Triggered AWS Glue jobs to clean data and query it via Athena and Redshift.
  - \* Created dashboards to visualize insights from processed data.
- **GitHub:** <https://github.com/trgtanh04/Mobile-AWS-Pipeline-Engineering>

### Movie Recommender System

December 2024 – Present

- **Key Technologies:** Python, Matplotlib, Scikit-learn, Streamlit
- **Description:**
  - \* Cleaned and analyzed movie data from Kaggle.
  - \* Extracted features and applied cosine similarity for movie matching.
  - \* Implemented a recommendation system using collaborative filtering.
  - \* Developed a Streamlit web app for personalized recommendations.
- **GitHub:** <https://github.com/trgtanh04/Movie-Recommendation-System>