Tran Duc Trung

Email: trung1803lucky@gmail.com
Phone: +(84) 0325647395
Github: TrungNotHot | Linkedin: Duc-Trung Tran | Facebook: Đức Trung Trần

1. Education

VNU-HCM University of Science

2021 – 2025 (Expected Graduation)

- Bachelor of Data Science
- GPA: 3.4

2. Skills & Coursework

- **Programming Languages:** Python, C/C++, R, Matlab, Java.
- **My tech stack:** Numpy, Pandas, Streamlit, Scikit-learn, Matplotlib, Pytorch, Spark, Dagster, dbt, MinIO, Docker, Linux (Ubuntu).
- **Database**: Microsoft SQL Server, PostgreSQL, MySQL.
- **Revelant Coursework:** Data Structures and Algorithms, OOP, Probability and Statistics, Discrete Math, Databases, Database Management Systems, Intro to AI, Data Mining, Machine Learning, Pattern Recognition.
- **Languages:** English (intermediate).

3. Projects

ETL Data Pipeline For NYC Trip Record

Oct – Dec. 2023

An ELT (Extract - Transform - Load) data pipeline with the TLC Trip Record Data

- Github: NYC-TripRecord
- Tasks: Develop end-to-end data platform following Lambda architecture. Build ETL pipelines, batch processing with Apache Spark.
- Serving: Analyze taxi journey details, including pickup and dropoff locations, and visualize them using Streamlit.
- Technologies: Docker, Dagster, Apache Spark, MySQL, MinIO, PostgreSQL, Streamlit.

House Price Prediction

May – June. 2023

- Github: House-Price-Prediction
- Tasks:
  - Using beautifulSoup and regex to crawl data on web
  - EDA, cleaning and preprocessing data (using MICE, one-hot encoding, scaling, remove outlier)
  - Build and compare to choose best model for predict house price (Linear Regression, Ridge Regression, Lasso Regression, Decision Tree, Random Forest, CatBoost, XGBoost, Stacking model)
- Technologies: beautifulSoup, regex, scikit-learn, matplotlib, streamlit

Searching Algorithms In Graph

Nov - Dec. 2022

A project using knowledge of DSA search algorithm in graph

- Github: Searching Algorithms
- Demo: Demo Searching Algorithms In Graph
- Tasks: Using pygame to create an interface, then applying Uniform Cost Search (UCS), Depth-First Search (DFS), and Breadth-First Search (BFS) algorithms to find a path between two nodes, while also leveraging Object-Oriented Programming (OOP) knowledge to aid in implementation.

4. Activities & Certifications

- Fundamental Data Engineering at #AIDE Institute
- Google Cloud Skills Boost #QuanQuanGCP
- HackerRank SQL (Basic to Advanced) Skills Certifications #HackerRank