



# HUNG LE TRI


DATA ENGINEER




---

 25/11/1997

 hunglt2511@gmail.com

 320 Khuong Dinh, Thanh Xuan, Ha Noi

 0989065487

 [www.linkedin.com/in/tri-hung-le](https://www.linkedin.com/in/tri-hung-le)

 <https://github.com/HungLe2511>

---

## INTRODUCE

I am a Data Engineer, with a passion and extensive knowledge of data science and technology.

During my career, I have had the opportunity to work with leading technologies and tools such as Hadoop, Spark(PySpark), Airflow and SQL.

I am learning and improving my system building skills but still have the ability to perform optimizations on my information using advanced techniques such as database tuning, query optimization, and balanced development load.

In addition, I am also particularly interested in researching and applying new technologies to improve the scalability and operability of system data.

With a eagerness to learn and the ability to work effectively in an environment requiring creativity and teamwork, with experience in managing small groups and working independently. I hope to be able to contribute to the project and give you advice for the development of the organization.

---

## WORK EXPERIENCE

**BKAV**

01/2021 - 09/2021

### Developer

project :

- (03/2021 - 06/2021) : Embedded engineer
  - Code to add features to smarthome devices (Electrical equipment system using zigbee)
  - AI Cam Operation Planning
  - API code for communication between Cam AI (in the provinces) and SQLite to save logs at Hanoi server using HTTP
  - Check the stable operation of Cam AI
- (06/2021 - 09/2021) : Training model for Cam AI facial recognition problem (recognizing faces with masks):

**Data Science**

project :

- Collect and clean data for tarot card project (using selenium for data collection and cleaning with pandas)
- Apply problems about the relationship between user and user, item vs item, user vs item, .. Visualize data with matplotlib. Code  
link: <https://colab.research.google.com/drive/1iEsubNcKbC59hBVq5SCWWJw3ovZNVmC?usp=sharing>
- Create datapine using spark to save to PostgreSQL
- Grasp and understand the process of collecting, processing and storing data
- Using docker and kubernetes to package test code that runs in another environment

**SAMSUNG SRV**

03/2022 - 09/2024

**Telecommunication system engineer**

project :

- (03/2022 - 05/2022): 2G, 4G, 5G knowledge training
- (06/2022 - 09/2024): Set up 2G GSM system equipment (Set up environment, internal network system, because it is Samsung's internal network, it is quite complicated and time-consuming):
  - Install operating system (Linux) and configure parameters
  - Update software (2-3 weeks/time) for system performance testing features in a simulated environment using E500 (add call stability feature, integrate 4G, 5G)
  - Coordinate with Indian Dev team to fix small errors encountered (working directly, I lead a small team of 3 people. Coordinate testing work and discuss with Indian Dev)
- (08/2023 - 05/2024): Setup SakTel 4G system equipment:
  - Install operating system (Linux) and configure parameters
  - Coordinate with Indian Dev team to fix minor bugs encountered

**Cty cổ phần chứng khoán DSC**

05/2024 - Hiện tại

**Data engineer**

project :

- Build a data warehouse to centralize stock trading data to support historical data analysis and track stock performance:
  - Perform data modelling for 2 broker and stock code. Helps synchronize the reading and writing process as well as assign tasks to each user that can affect the warehouse.
  - Tech use : Source data from postgres, MySQL, PostgreSQL for warehouse, spark for ETL, deploy using docker
- Maintain and make new pipeline for get data from stock trading APIs, data cleaning pipeline, stock index calculation pipeline, broker behavior pipeline :
  - Perform data, feature data, prioritize data flow, make datamart for each end user.
  - Tech use : Spark, docker

**EDUCATION****HaNoi University of science and technology**

2016 - 2021

**Electronics and Telecommunication**

CPA : 2.8

## SKILL AND PROJECT

Database, storage : Postgresql , S3

Framework : Spark, airflow, hadoop (basic),kafka , pandas, seaborn, Power BI, tensorflow.

Language : Ability to read and understand documents, basic communication.

Additional skills: git, github, docker,k8s.

OS : linux, centos7

Course learning :DeepLearning (on youtube of Andrew Ng), Cloud Hadoop: Scaling Apache Spark ,

Advance Your Data Skills in Apache Spark , Apache Spark Essential Training: Big Data Engineering , Stream Processing Design Patterns with Spark.

Project:

- PySpark for Beginner

Url : <https://github.com/HungLe2511/PySpark-for-Beginner>

Framework : Spark, airflow, AWS S3, postgresSQL.

- Crawl stock data streaming kafka, warehouse and lake:

URL : [https://github.com/HungLe2511/ETL\\_MySQL\\_minio\\_postgresQL\\_Docker](https://github.com/HungLe2511/ETL_MySQL_minio_postgresQL_Docker)

Bank prj :

Make data pipeline with some tools

Tools : storage(Oracle, MinIO, document), process(Spark, dbt), visualize(Power BI), Airflow

Link pipeline : [https://lucid.app/lucidchart/adba108-9102-41c2-b0d4-be8da291bf46/edit?](https://lucid.app/lucidchart/adba108-9102-41c2-b0d4-be8da291bf46/edit?viewport_loc=-235%2C205%2C2136%2C1140%2C0_0&invitationId=inv_b517dedf-145c-4fc2-93f6-2423eec884fa)

[viewport\\_loc=-235%2C205%2C2136%2C1140%2C0\\_0&invitationId=inv\\_b517dedf-145c-4fc2-93f6-2423eec884fa](https://lucid.app/lucidchart/adba108-9102-41c2-b0d4-be8da291bf46/edit?viewport_loc=-235%2C205%2C2136%2C1140%2C0_0&invitationId=inv_b517dedf-145c-4fc2-93f6-2423eec884fa)

---

## CERTIFICATE

**Aptis Certificate (2021)**

**B2 (HaNoi University of science and technology)**

**SQL (Advanced) Certificate (2023)**

<https://www.hackerrank.com/certificates/402b32cf94ec>

**Python (Advanced) Certificate) (2023)**

<https://www.hackerrank.com/certificates/de5692831028>

---

## PRIZE

**Outstanding Employee of the 2nd Quarter**

04/2023

Excellent

SAMSUNG SRV