Dao Van Thang
DATA ENGINEER

📅 Jan 27,2000 👤 Male

📞 0394927609 ✉ thangdao2701captain@gmail.com

📍 Ho Chi Minh , Viet Nam 🔗 www.linkedin.com/in/thangdaoda taengineer

OBJECTIVE

I am a sociable person and can work hard to absorb new knowledge, I want to become a Senior Data Engineer 3 - 4 years next.

I am confident in my ability to teamwork, handle situations as well as high-pressure resistance. My English in communication is pretty good.

ADDITIONAL INFORMATION

- At the moment, I am an Data Engineer. I have good knowledge of Python, Java, Cloud System, Airflow, Hadoop, Spark, Bigquery, Datawarehouse, DataLake, Data modeling, Hive, Sqoop, MySQL, SQL Server. I want to work in an environment that I can work and learn more about Data Engineer and how to create benefits for an organization. I have a toeic score of 580.

EDUCATION

HCMC University of Technology and Education

(Aug 2018 > June 2022)

Major: Data Engineer

GPA: 2.89/4

WORK EXPERIENCE

Amanotes PLD Company (Nov 2021 > Present)

Data Engineer - full-time

Main responsibilities:

- Building ETL and Data Warehouse. Data pipeline from many sources like MongoDB, Bigquerry, App event, API .....Ingestion Data
- Using google cloud platform, Set up Airflow composer, DAGs ,tasks . Set docker for dev environment. Use dbt for QC data.
- Maintain data pipeline from app event to data warehouse for business .

Recognition and skills gained:

- Cloud System and build Airflow composer.
- Building data warehouse and data modeling.
- Exp about a pipeline for many sources data.
- Improve Communication Skill and team working.

Officience Company (June 2021 > Oct 2021)

Data Engineer - full-time

Main responsibilities:

- Scrap data, crawling, parsing Html, import data to DBMS, report data downloaded every month.
- Deploy ETL. Assist to control an HDFS through HIVE table. Load

data in HDFS by Spark RDD. Transfer data by Spark.
- Import data from Mysql to HIVE by apache Sqoop and vice versa.
- Deploy hadoop in docker.

Recognition and skills gained:
- Report data and do tasks on time.
- Improve use git, linux, bash script.
- Learn more about crawling data, build a bigdata system.
- Improve English Communication Skill.

SKILLS

Skill	Level
Python	●●●●○
Java	●●●○○
Scala	●●○○○
Data Modeling	●●●○○
Cloud System	●●○○○
SQL/RDBMS	●●○○○
Spark and Hadoop	●●●○○
Docker	●●●○○
ETL	●●●○○
English	●●●○○

CERTIFICATIONS

Certificate of merit from the Ho Chi Minh City Student Union in the volunteer movement 2020.
(2020)

PERSONAL PROJECTS

DATA WAREHOUSE: Building a data pipeline with data crawling (Oct 2021 > Jan 2022)
Full stack
Usage: Python, Scrapy, Hive, Spark, Apache superset .
Function: Crawling data, Cleansing data, ETL, Visualize data in data warehouse.
Experience: Python, Scrapy, Data Warehouse, Apache superset …
GIT: github.com/captionnemo/Datapipeline_DataEngineer_project_2022

HADOOP: Use MAP-REDUCE with HADOOP
 (Mar 2021 > Apr 2021)
Full stack
Usage: HDFS, Java MapReduce, Ubuntu Shell Script.
Function: Format, Reformat, Multiply, Count number Word in data.txt.
Experience: HDFS, Java MapReduce, Ubuntu Shell Script, …
Git: github.com/captionnemo/CountWord-HadoopMapreduce.git

DOCKER: Build Hadoop, Spark, Hive, Pig, Sqoop on docker composer   (July 2021 > Sep 2021)
Full stack
Usage: HDFS, Spark, Hive, Pig, Apache Sqoop, Dockerfile and dockercompose, Shell Script.
Function: Build a Bigdata system by docker-compose.xml
Experience: Dockerfile, docker-compose,HIVE config, Sqoop config.
Git: github.com/captionnemo/hadoop-Spark-Hive-Pig-Sqoop_on_docker.git

BIGDATA ANALYST: PREDICT CANCER USE SPARK

 (Apr 2021 > June 2021)

Full stack

Usage: python, spark SQL,RDD , HDFS, Machine Learning.
Function: Train data with RandomForest Algorithm,Predicting Probability of Malignant and Benign BreastCancer with Random Forest Model.
Experience: Python, Spark SQL, RDD, HDFS,RandomForest models....
GIT: github.com/captionnemo/SPARK-BIGDATA-ANALYST-PREDICT-CANCER.git

JAVA PROJECT: HELMETS COMPANY MANAGER WEBSITE

 (Apr 2020 > June 2020)

Full stack

Usage: Java Servlet, SQL Server, JSP, Bootstrap, Java Script.
Function: Manage employee,task, project, product, salary
Experience: Java Servlet,API , SQL Server, Java Script,…
GIT: github.com/captionnemo/Helmet-Company.git