



CANDIDATE REPORT

Position: **Data Engineer**

DETAILS

Name: __ Nguyen	YOB: 19__
Gender: Male	

OVERALL ASSESSMENT

Mr. __ has 5 years of experience working in software development as a Data Engineer. He has hands-on experience in __ and __. He is strongly skilled in optimizing existing flows for better speed and higher stability; processing data and gaining insights from them to generate recommendations. Besides technical skills, Mr. __ possesses a strong desire to expand his current skillset in a fast-paced work environment. His English is at an advanced level so he confidently communicates in English. In addition, Mr. __ is professionally responsible and careful in work, able to work independently, self-directed, and with teamwork.

EDUCATION & CERTIFICATES

Hanoi Foreign Trade University 2010 - 2014
Bachelor of Economics and International Business

SKILLS

Technical skills:

Spark	Python	FastAPI	Redis
RabbitMQ	Airflow	MySQL	Docker
Flask	MongoDB	PySpark	AWS Glue
Redshift	Azure K8s	Tableau	PostgreSQL
Hadoop	Netezza	Spark Scala	Scala
SQL Server	SSIS		

Soft skills:

- Can study new technology quickly
- Able to work under minimal supervision, adhere to deadlines
- Able to rotate concentration smoothly and entirely from one task to another. Active observation and listening team members.
- Devoted to supporting people when they need it.
- Can work independently and with teamwork
- Good Communication and Relationship Development
- Good problem-solving, analytical, and critical thinking skills
- Good English (4 skills)



WORKING EXPERIENCE

__ JSC Sep 2021 – present *Senior Data Engineer*

__ JSC: __ is a trusted IT outsourcing company in Vietnam, with 1000+ headcounts and over 40 projects currently running.

❖ **Project: __ Financial Group**

• __ Group is an online trading provider, offering impeccable execution and trading conditions including very low spreads, professional services, dedicated support and powerful tools.

• **Description:**

- Build Restful API to support new trading platform app
- Build migration ETL to migrate data to new database

• **Technologies:** Python, FastAPI, Redis, RabbitMQ, Airflow, MySQL, Docker

• **Responsibilities:**

- Developed RESTful APIs that serve users requests, publish message to RabbitMQ and then receive message to trigger other APIs or to save changes into DB.
- Developed Migration jobs via Airflow to migrate data into new DB.
- Restructured the project so that it is easier to manage, develop and maintain

❖ **Project: __**

• ____ is a subsidiary of ____, which is the only Fortune 500 company in Malaysia.

• **Description:**

- Build ETL jobs to pull data from several sources, including MongoDB, CDP system and S3 buckets, and then process data and dump into S3 bucket.

• **Technologies:** Python, Flask, MongoDB, MySQL, Docker, PySpark, AWS Glue, AWS Lambda, Redshift

• **Responsibilities:**

- Worked with ETL jobs that deal with data of tens of millions of records
 - Optimized MySQL database by setting indexes and optimized SQL queries against millions of records to reduce query time from 30 minutes to under 2 minutes.
 - Developed modules in RESTful API system that pulls data from several sources such as MongoDB, CDP system, S3 and then processes data and dumps them into S3.
 - Developed PySpark jobs in AWS Glue and AWS Lambda to run data pipelines among several sources like S3, Redshift

❖ **Project: ____ Pharma**

• ____ Pharma is one of the largest healthcare services groups in Asia.

• **Description:**

- Build data flow to send reports daily or monthly and optimize dashboard to load faster



- **Technologies:** Python, Docker, Azure K8s, Tableau, PostgreSQL
- **Responsibilities:**
 - Optimized old scripts to automatically pull data and send weekly or monthly reports to customers. Packaged these scripts into Docker images and deployed these images to Kubernetes on Azure cloud. Optimized version ran much faster, up to 10 times than the previous version, and reduced RAM usage from gigabytes to small footprint.
 - Optimized existing Tableau dashboard so that the report loading speed was much faster than previous version, around 2-3 times faster.
 - Created Tableau dashboard that detects suspicious scan activities over times. This helps clients notify any unusual activities and make necessary investigations.

Bank _____ of Vietnam Nov 2019 – Aug 2021

Data Engineer

❖ **Analytics Team – _____ Center**

- _____ is one of the largest bank in Vietnam by total assets.
- **Technologies:** Python, PySpark, Hadoop, Netezza
- **Responsibilities:**
 - Worked with data of over 10 million users with tens of millions of transactions per month.
 - Researched the data systems of the bank, transferred important data to Sandbox for the Data Analytics Team to analyzed and processed.
 - Managed and maintained two workstations and installed Spark with Yarn on them so that the Data Analytics Team could use it for analyzing big data. All components were installed manually by hand.
 - Cooperated with IT Center to set up the Customer Personas System by designing Analytics area based on Hadoop system and transferring historical customer data into that area.
 - Cooperated with Card Center to deal with some use cases about credit card, such as managing credit limit and predicting which product or service category that customers are likely to purchase next.
 - Built Customer Churn model to predict the probability that customers will no longer use SmartBanking service. The AUC score is high as 0.91

Viettel Cyberspace Center Nov 2018 – Oct 2019

Big data analyst

❖ **_____ Center**

- _____ is one of the largest telecommunications service provider in Vietnam.
- **Technologies:** Python, Spark Scala, Scala, MySQL
- **Responsibilities:**
 - Worked with data of over 100 million users with hundreds of millions of transactions per day.



Reco

Connect Our World

- Analyzed, audited and processed Telco data to add features to User Profile system, which contains many features about users such as gender, age, location and so on. This system will be used to build machine learning model for many purposes in the future.
- Responsible for target marketing in accident insurance product. Mined telco data logs to find key features, built machine learning model to score users and used statistical methods to generate daily lead to send marketing SMS to potential customers.

Improved the conversion rate by up to 6 times as compared to mass marketing

_____ Bank Feb 2017 – Oct 2018

Data analyst

- _____ is one of top 5 largest banks in Vietnam by total assets.
- **Technologies:** Python, SQL Server, SSIS
- **Responsibilities:**
 - Worked with data of over 5 million users with millions of transactions per day.
 - Processed data and created daily report to monitor the results and performance of loan and deposit of business units of the bank in accordance with BOD strategic plan.
 - Prepared monthly reports of loan and deposit to ALCO (BOD) Meeting for developing strategic plan of the bank.
 - Planned the monthly averages of loans and the lending interest rate of business units to ensure the Net Interest Income (NII) target.
 - Applied model to estimate the fluctuation in NII when changing the deposit interest rate (accuracy over 60%).
 - Restructured data and improved reports to reduce the processing time by 30%.