Resume
Van-Duyet LE

Ho Chi Minh City | Year of birth: 1995
w: https://duyet.net | e: me@duyet.net
Github: https://github.com/duyetdev | Linkedin:
https://linkedin.com/in/duyet

I am a Data Engineer. Always bet on #Machine_Learning and #Javascript.
Currently, I'm working on Data Engineering (Data Pipeline, DWH, Data Processing and Big Data based on AWS and GCP), Machine Learning (Classification, Clustering, ...) especially NLP (Sentiment Analysis, Text Classification, Word2vec, ...).

Experiences

OCTOBER 2018 - CURRENT

Data Engineer – *Fossil Group Inc*

Working on Data Pipeline and Data Warehouse project

- Build a Data Pipeline to process big data in AWS and GCP. Technologies: AWS (K8S, S3, EMR, Lambda), Apache Airflow.

JUNE 2017 - OCTOBER 2018

Senior Developer – *FPT Software*

Working on Recommendation System and Data Lake project

- Build a Data Pipeline to process 2TB of data per day in AWS. Technology: AWS (EC2, S3, EMR, CodePipeline, Lambda), ElasticSearch, Python, PySpark.
- CI/CD for automatic deployment (CodePipeline, CodeBuild, Docker, ECS).
- Data Pipeline on Microsoft Azure (Data Lake, Data Factory, Azure Functions, Databricks).

SEPTEMBER 2016 - JUNE 2017

Data Engineer – *John von Neumann Institute - Vietnam National University*

Main contribution: Team member of two projects:

- Headhunter Company
    - Cleaning, building data pipeline and ML model for Candidate matching.
    - Visualization dashboard: hiring competitiveness, ideal candidates, etc.
    - Project: Skill2vec - Relevant Skills from Job Description: Skill2vec is a neural network architecture inspired by Word2vec - developed by Mikolov et al. It transforms skills to new vector space, which has the characteristics of calculation and presents skills relationships (https://arxiv.org/pdf/1707.09751.pdf).
    - Technology: D3.js, Angular.js, Python, ElasticSearch, Cloudera.
- Advertising Network Company
    - Predict the Demographics of users from tracking data (Movie ads network).
    - Technologies: Apache Spark (PySpark), Hortonworks.

OCTOBER 2014 - SEPTEMBER 2016

Full-stack Developer – *Information System Lab, UIT*

Main contribution: Working on many projects such as:

- Self-serve ad platform for marketing, Real-time Bidding server by Nodejs.
- Web Spider, Crawler. Research and Develop websites with Node.js/Javascript. Real-time dashboard with D3.js, Node.js, Kafka, and Socket.io.

Skills

Web Development: Javascript & Node.js (Angularjs, D3js, Reactjs), Python (Flask, Django), SQL/NoSQL.

Data Engineer: AWS (EC2, ECS, EMR, S3, CodePipeline), MS Azure (Azure Function, Data Factory, Databricks), Python (flask, pandas, scrapy, scikit-learn, NLTK, ..), Javascript (natural.js, Phantom.js, D3js, Vuejs, Reactjs), Apache Spark, Apache Airflow.

Education

JULY 2018 - CURRENT

Master of Computer Science | **University of Information Technology - HCMC**

JUNE 2013 - 2017

Bachelor of Information System | **University of Information Technology - HCMC**
Thesis: *Build a network of career skills and support optimal job search.*

Awards

Third Prize "Entropy 2016" – *John von Neumann Institute (VNU).*

The first competition of Data Science in Vietnam ⬈

Third Prize "FOSS 2015" – *Hanoi University of Business and Technology.*
Free and open-source software (FOSS) competition.

Final round "Creative IT Camp 2013" – *University of Information Technology (VNU).*
Innovation and Programming hackathon competition.

References

MEng. **An DANG** – *Tech Lead of Glandore System Project*
John von Neumann Institute (VNU) | email: an.dang@jvn.edu.vn

MSc. **Thinh PHAM**
John von Neumann Institute (VNU) | email: thinh.pham@jvn.edu.vn