

**DANG CHIEN NGUYEN**

Hanoi, Vietnam

☎ +84383495684 ✉ dangchienhsgs@gmail.com

WORK EXPERIENCE**ZALO Group**

5/2017 - 4/2018

Senior Lead Data Scientist (4/2020 - Present)

Responsible to design architecture and maintain a recommendation system for Zalo, the most popular OTT in Vietnam with more than 100 million users. I lead a full-stack data science team that does both ETL, modeling, and serving our models to guarantee the consistency and the balance between researching and deployments in our product.

Data Engineer related:

- Develop a high-performance in-memory graph service that provides the best state-of-the-art implementations of graph algorithms (PageRank, Community Detection, Graph Embedding). These services and data structures are the core of many intelligent features and operations at Zalo like Friend Suggestion, Smart Notification, Group Suggestion, serve personalized-per-user suggestions for more than 50 million users active daily effectively.
- Develop and maintain ETL pipelines for feature engineering, benchmarking, monitoring the result of production in Pyspark, Hadoop, Airflow.
- Build high-performance workers and services to process both real-time events and on-scheduled events. Our system adapts and provides up-to-date predictions base on user behaviors. The system was mainly built on Java, RabbitMQ, Thrift, and some C++ libraries for high-performance purposes.
- Build a Machine Learning process and ML Ops such as Model iterations, Model Quality Checking, and Data Quality systems to reach both quality and quantity KPI by optimizing many robust and explainability models, ensuring our product health and data health.

Data Science related:

- Track record of publications and contributions to the machine learning community at large, especially in Social Network and Knowledge Graph to choose the best suitable solutions for our products, keep the balance between the complexity, the cost, and the scalability and the accuracy.
- Responsible for the direction, development, and vision of the team. I am leading 6 researchers to implement, deploy, and analyst the feedbacks to continuously improve our models by iterations, help the product reach both Tier-1 and Tier-2 KPIs, using many techniques like CatBoost, Xgboost, Stacking, Ensemble, Categorical Encoding, LDA,...
- Working with Data Analyst teams and Product Vertical teams to define the important metrics and KPIs. I must learn deeply about the product to keep the best direction and importance of our team to the products.

ZALO Group

4/2019 - 4/2020

*Lead Data Scientist***ZALO Group**

9/2017 - 4/2019

Senior Data Scientist

Ecomobi PTE <i>Data Mining Developer</i> <ul style="list-style-type: none">- Design ETL and Prediction pipelines in python, pandas and scikit-learn, xgboost and docker.- Manage Hadoop cluster and design data flow based on Hadoop, Spark, ElasticSearch to analyze 60 million Vietnam Facebook profiles and get useful information to help lead scoring system better	9/2016 - 5/2017
---	-----------------

VCCorp <i>Big Data developer</i> <ul style="list-style-type: none">- Working with high volume data system using Big Data Technology Stack: Hadoop, Spark, Hbase,..- Design a low latency protocol based on redis to improve the communication between servers- Work standard-alone for building A/B testing framework in Java in both server and client.	4/2016 - 9/2016
--	-----------------

Data Science Laboratory (Hanoi University of Science and Technology) <i>Researcher</i> <ul style="list-style-type: none">- Focus on the graphical model and its application in classification and clustering problems, especially for text data.- Working on programming and benchmark many Machine Learning algorithms in Python and C++.- Published an article for Knowledge And Information System (KAIS) in an HDP (Hyber dynamic parameters) algorithm. (See [1] at References).	3/2013 - 7/2017
---	-----------------

EDUCATION

Hanoi UNIVERSITY OF SCIENCE AND TECHNOLOGY <i>Center for Training of Excellent Students (CTES)</i> Information Technology	Sep 2012 - JULY 2017
--	----------------------

HIGH SCHOOL FOR GIFTED STUDENTS, VIETNAM NATIONAL UNIVERSITY <i>Class of Mathematical</i>	Aug 2009 - 2012
---	-----------------

CERTIFICATIONS

Honor title Second Prize of Award of “Student’s Scientific Research” 2016, Hanoi University of Science and Technology	2015
---	------

PUBLICATION

- Have an accepted article in Knowledge and Information System journal (KAIS) about discovering new classification method in Hyper Dynamic Parameter (HDP) at (1)	2016
---	------

TECHNICAL SKILLS

Experienced Data Scientist with a demonstrated history of working in the internet industry. Skilled in Apache Spark, Big Data, Python (Programming Language), Data Engineering, and Hadoop. Strong in both applied science and data engineering.

Specialties:

- * Proficient with Java, Python, familiar with C++.
- * Fast learning and critical thinking. Language is not a problem for me.
- * Knowledge in Machine Learning and Big Data Computing both in academics and production are my strong points.

REFERENCES

1. My journal in Springer: [http://link.springer.com/article/10.1007/s10115-016-0956-6?](http://link.springer.com/article/10.1007/s10115-016-0956-6?wt_mc=Internal.Event.1.SEM.ArticleAuthorAssignedToIssue)
wt_mc=Internal.Event.1.SEM.ArticleAuthorAssignedToIssue

