QUAN DUONG

Data Scientist

📞 +84938134558   @ quanduongbao@gmail.com   🔗 https://quandb.github.io

EXPERIENCE

Data Scientist

Tenpoint7

📅 09/2016 - Ongoing   📍 Ho Chi Minh

- Build analytic solution that solve variety of use case for customers
- Develop more than 5 information retrieval features for data product of company
- Work closely with Data Engineering and Full Stack Development team to deploy Machine Learning model on the cloud

Software Engineer

Infonam

📅 04/2015 - 10/2015   📍 Ho Chi Minh

Responsible for automation testing of new released features of network devices

Software Engineer

ISePRO

📅 02/2014 - 04/2015   📍 Ho Chi Minh

Audit security level of network devices (routers, switches, firewalls)

EDUCATION

Master of Computer Science

University of Science - Vietnam National University

📅 09/2015 - 01/2019   📍 Ho Chi Minh

Bachelor of Information Technology

University of Science - Vietnam National University

📅 09/2010 - 07/2014   📍 Ho Chi Minh

CERTIFICATES

Statistical Learning

Stanford Online

Structuring Machine Learning Projects

Coursera

SKILLS

Machine Learning

Feature Selection

Python

Text Mining

Ensemble Learning

Research and Development

Leadership

Computer Network

LANGUAGES

Vietnamese
Native

English
Advanced

PROJECTS

Automatic labeling for topic models

Develop company product's features

- Employed state of the art method for automatic labeling of topics in topic model from performing the literature review and experimenting several approaches with multiple data-sets
- Work closely with Data Engineer and Full Stack Development team to deploy the machine learning model to production

Automated Trajectory Clustering

📅 01/2018 - 11/2018
🔗 https://github.com/quandb/atc

Support to detect the typical air routes between two airports based on ADS-B data. Specifically, the flight traffic will be classified into major groups based on similarity measures, which helps to reduce the number of flight paths between airports.

Staff Rostering Optimization

📅 06/2017 - 08/2017

The initial version of roster engine reduce the cost 6.7% in total, and up to 26.3% for Over Time case.

Customer Churn Analysis

📅 05/2017 - 06/2017

- Employ visual analysis to extract the insight from human interaction
- Hepl client to do identify the factors that influence to customer churn and build predictive model to predict the churn

Network Graph Analysis

📅 10/2016 - 12/2016

- Detect the influencer in the network graph
- Predict the employee churn based on their interactions
- Implement http api to provide an interface for querying from outside

Sentiment Classification

📅 09/2016 - 10/2016

Working on improving sentiment classification accuracy and collaborating with stakeholders to bring machine leaning model to production

Text Classification

📅 08/2016 - 09/2016

- Help the client to improve the machine learning model performance in term of accuracy and time
- Dealing with multi-class classification (15 classes), imbalanced data

TECHNOLOGIES

Data Science

`sklearn` `xgboost` `pandas`
`spark` `matplotlib` `tableau`
`nltk` `git` `h2o` `mysql`
`mongodb` `numpy` `scipy`
`pattern` `aws` `jira` `intellij`
`gensim` `networkx` `linux`

REFERENCES

Dr. An Mai
an.mai@jvn.edu.vn

PUBLICATIONS

A Simplified Framework for Air Route Clustering Based on ADS-B Data

The 2019 IEEE-RIVF International Conference on Computing and Communication Technology

Quan Duong, Tan Tran, Duc-Thinh Pham, An Mai

📅 03/2019   🔗 https://bit.ly/2VoSVtS

An Adaptive Hash-Based Text Deduplication for ADS-B Data-Dependent Trajectory Clustering Problem

The 2019 IEEE-RIVF International Conference on Computing and Communication Technology

Tan Tran, Quan Duong, Duc-Thinh Pham, An Mai

📅 03/2019   🔗 https://bit.ly/2EN4cis