

HOANG-THONG VO

☎ (+84) 522004xxx | 🌐 [thongvhoang](#) | [in](#) [thongvhoang](#) | [k](#) [thongvhoang](#) | ✉ [vhthongvn](#) | 🌟 [hoangthongvo](#)
| [GoogleScholar](#) Hoang-Thong Vo

RESEARCH INTERESTS

My current research interests focus on three topics: intelligent transportation systems, smart mobility, and computer vision for smart cities.

EDUCATION

- **VNUHCM - University of Information Technology** *Aug 2018 - Aug 2022 (Expected)*
Bachelor of Science, major: Data Science, in-major GPA: 8.76/10.00 (3rd year - 5th Semester)
Coursework: Statistical Machine Learning, Deep Learning, Optimization and Applications, Data Analysis and Visualization, Experimental Design Analysis, Data Collection and Preprocessing, Data Structure and Algorithm, Object Oriented Programming, Probability and Statistics
- **Le Quy Don Khanh Hoa High School for the Gifted** *Aug 2015 - May 2018*
High school education, class of Informatics

WORK EXPERIENCE

- **Spring Knowledge Global Joint Stock Company** Apr 2021 - Current
Data Scientist
 - Collected and built a dataset to train model using **Labelimg** open source and **Roboflow** app
 - Researched and developed a model to tackle the small object like QR-code in object detection using **Tensorflow** and **Pytorch** framework
 - Deployed PyTorch model on Android app with **Tensorflow Lite**
- **Officience** Aug - Oct 2020
Machine Learning Engineer
 - Crawled, optimized crawling process and pre-processed data from 5 real estate sites by **Scrapy**, **beautifulsoup4**, **lxml**
 - Performed **Bash** script that automatically inserts raw data into databases
 - Cleaned and wrangled raw data on a database using **Pandas**
 - Visualized data using **Matplotlib**, **Seaborn** and **PySpark**

RESEARCH EXPERIENCE

- **Information Systems Laboratory - University of Information Technology** Jul 2020 - Current
Research Assistant
 - Supervisor: Dr. Trong-Hop Do
 - Research on Visible Light Communication, Vehicle Communication and Sensing, and Optical Camera Communication

PUBLICATIONS

- **Hoang-Thong Vo**, Ngan-Linh Nguyen, Thien-Long Nguyen, Dinh-Tu Vo, and Trong-Hop Do, “Xây Dựng Bộ Dữ Liệu UIT-DET Cho Hệ Thống Giám Sát Giao Thông Trong Thành Phố Thông Minh”, 2021 10th CITA National Conference on Information Technology and Its Applications (**Accepted**)
- Trong-Hop Do, Ngan-Linh Nguyen, **Hoang-Thong Vo**, Thanh-Binh Nguyen and N. Tan Vu Khanh, "Deep learning based image processing for proactive data collecting system for autonomous vehicle," 2021 21st ACIS International Winter Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD-Winter), 2021, pp. 253-256, doi: 10.1109/SNPDWinter52325.2021.00062.

PROJECTS

- **Predictive models for diabetes mellitus** *Feb 2021*
Description:
 - A personal side project experimenting on multiple binary labels tabular data classification models.
 - Ranked in the top **47% percentile** of a data science competition on **Kaggle** with accuracy more than **85%** and the highest accuracy on that competition was over 87%**Techstack:** Scikit-learn, Numpy, Pandas, XGBoost
- **Classification of food for holidays and special occasions on the Epicurious - Recipes with Rating and Nutrition dataset** *Nov - Dec 2020*
Description: Exploratory data analysis such as heat map and descriptive data statistics, handled missing data, experimented methods in feature scaling (Robust Scaler, Standard Scaler, ...), tried with classification models and deployed web app machine learning models
Role: Team Leader
Techstack: Scikit-learn, Matplotlib/Seaborn, Numpy, Pandas, Streamlit, Heroku
- **Analyzing the effect of these factors on students' final grades** *Jun - Jul 2020*
Description: Exploratory data analysis, analysis of variance factors affecting final grades, regression analysis (Linear Regression, Support Vector Regression)
Role: Team Leader
Techstack: R, Python, Scikit-learn, Pandas
- **House numbers recognition on The Street View House Numbers Dataset using Convolutional Neural Networks** *Jun - Jul 2020*
Description: Built the model which is able to recognize digits and numbers in natural scene images. The model was built on 2 dataset formats with 94% and 97% accuracy, respectively evaluated by the accuracy metrics.
Role: Team Leader
Techstack: Scikit-learn, Keras, OpenCV
- **Web Scraping** *Apr 2020*
Description: Virtualized the Covid19 data from the Worldometer with InfluxDB. The model describes the total cases, daily new deaths for each country in the world.
Techstack: Scrapy, Python, InfluxDB
- **Artificial Intelligence Nutritionist** *Sep - Dec 2017*

Description: Diet Recommendation System for Healthcare Service Based on User Information using Neural Network in Deep Learning and FatSecret Food Database. This project awarded Encouragement Prize in the Science and Technology Contest for High School Students, Khanh Hoa Province, December 2017.

Techstack: Python, Keras, Numpy, FatSecret Platform API

ACCOMPLISHMENTS AND ACTIVITIES

- **Half Marathon finisher at Techcombank Ho Chi Minh City International Marathon** *Apr 2021*
Ran 21.1 km around Ho Chi Minh City, Vietnam. The race was organized with the participation of 10,706 runners at distances of 5, 10, 21, and 42 km.
- **Member of Free Contest** *Feb 2016 - Feb 2021*
These contests are for young coders in different regions in Vietnam to improve their coding skills with our high-quality problems.
- **Participated in the SeABank Run for The Future annual race** *Nov - Dec 2020*
Athletes who complete the registered distance, sponsors will donate 1000 VND / 1km of valid runners. All funds raised by SeARun 2020 will be transferred to the "Dream Nurturing" fund and used to award at least 26 scholarships to poor studious students and young talents across the Vietnam country.
- **Participated in the VNU Will Run Project** *Nov - Dec 2020*
Athletes who complete the registered distance, sponsors will donate 1000 VND / 1km of valid runners. The "VNU Will Run" is an annual project organized by the Ho Chi Minh City National University Youth Union's Personnel Committee (VNU-HCM) organized to raise funds for the "Will Foundation" scholarship to support exemplary students who are energetic and rise in life.
- **Participated in the Online Running Race - Cashless Day organized by Saigon Thuong Tin Commercial Joint Stock Bank and Tuoi Tre newspaper** *Jun 2020*
The running race aims to develop the habit of non-cash payment & exercise habits to improve people's health. At the same time, for each athlete participating and completing the registration distance, Saigon Thuong Tin Commercial Joint Stock Bank and its partners will contribute 100,000 VND to support the frontline medical team in the fight against Covid-19.
- **Participated in the Shopee Code League 2020 (Student Category)** *Jun - Aug 2020*
Shopee Code League is a 3-week coding challenge consisting of 3 coding competitions open to all students and professionals across the region. The competitions, specially designed by the Shopee tech teams, cover data analytics, data science and algorithmic problems. Participants must analyse the datasets, draw insightful conclusions and solve the problems in a specified amount of time.
- **Honor prize in Provincial Informatics Olympiad for High School Students** *Nov 2016*
Competed individually in 2 days to solve 3 data structures and algorithms problems in 3 hours on each day.
- **Gold medal in 30/4 Informatics Olympiad for Vietnam Southern High School Students** *Apr 2016*
Competed individually in 1 day to solve 3 data structures and algorithms problems in 3 hours.

LANGUAGES

- **Vietnamese:** Mothertongue
- **English:** Reading, Writing: Intermediate; Listening, Speaking: Basic

TECHNICAL SKILLS

- **Programming Language:** Python, C/C++, Bash script, SQL (Postgre), R
- **Machine Learning Algorithms:**
 - Supervised: Regression (Linear, Logistic), Memory Instance-based (k-Nearest Neighbor), SVM, Decision Tree (CART), Ensemble (Bagging, Random Forest, Boosting), Naive Bayes

– Unsupervised: Clustering (K-Means), Linear Dimensionality Reduction (PCA, LDA)

- **Deep Learning Frameworks:** Scikit-learn, Keras, Tensorflow
- **Web Scraping Frameworks Proficient:** Scrapy, Request, BeautifulSoup, lxml
- **Data Analysis and Visualization Frameworks:** Pandas/Dask/Ray/Modin, Matplotlib/Seaborn, Scipy, Numpy
- **Version Control Proficiency:** Git, Github
- **Document Creation:** Microsoft Office Suite, LaTeX, Markdown

CERTIFICATIONS

- Computer Vision Basics - Coursera
- Fundamentals of Visualization with Tableau - Coursera

REFEREES

Dr. Trong-Hop Do	Senior Researcher at Faculty of Information Science and Engineering, University of Information Technology, Vietnam National University Ho Chi Minh City hopdt@uit.edu.vn
------------------	---