Huyen Thuc Khanh Vo

Ho Chi Minh, Vietnam
✆ *(+84) 947 786 682*
✉ *vothuckhanhhuyenvn@gmail.com*
🖥 *vothuckhanhhuyen.github.io*

Research Interests

Present My research focuses on enhancing the *performance*, *efficiency*, and *robustness* of **machine learning models** and quantify *uncertainty* in the predictions that these models produce by using two core approaches: (1) **optimization methods** (momentum and adaptive optimizers); (2) **statistical methods** (conformal prediction, bayesian and probabilistic).

Education

2019 - 2023 **Hanoi University of Science and Technology (HUST)**, *Hanoi, Vietnam.*
- B.Sc, Major in **Computer Science** (Honors Program), CPA: 3.87 / 4.0 (Class rank: 3/35).
- Thesis: Continuous Neural Ordinary Differential Equations [pdf] (**Best Thesis Presentation Award**). Supervised by Dr. Linh V. Ngo.

Experience

Nov 2021 - Present **AI Research Resident**, *FPT Software AI Center*, ai.fpt-software.com.
- Advisors: Prof. Tan M. Nguyen (NUS) & Dr. Thieu N. Vo (Ton Duc Thang University).
- Research topics: neural odes, conformal prediction and calibration, graph neural networks.
- Proposed novel methods to improve *performance* and *efficiency* of neural odes using optimization methods (nesterov accelerated gradient (momentum optimizer), rmsprop (adaptive optimizer)).
- Ran experiments of Nesterov Neural ODE, RMSprop Neural ODE on time series, continuous normalizing flows, image classification, and point cloud separation.
- **Ongoing project:** Quantified uncertainty of continual learning models using conformal prediction.
  - Ran experiments on class incremental learning settings to verify that continual learning models perform poorly on old tasks in terms of confidence coverage and prediction set size.
  - Proposed a novel method based on exchangeability property to address these limitations.

Apr 2021 - Nov 2021 **Undergraduate Research Student**, *Data Science Laboratory*, ds.soict.hust.edu.vn.
- Advisors: Dr. Linh V. Ngo (HUST) & Prof. Khoat Q. Than (HUST).
- Research topics: continual learning, probabilistic graphical models, variational inference.
- Explored four primary approaches in continual learning, analyzed the strengths and weaknesses of each approach, made comparisons and suggested some improvements.
- Re-implemented key algorithms from each approach.

Jul 2021 - Sep 2021 **Data Analyst Intern**, *VinBigdata*, vinbigdata.com.
- Advisors: Dr. Tham H. Hoang (University of Connecticut).
- Examined responses to the H3N2 influenza challenge using machine learning models.
  - Dataset: GSE61754, genome-wide gene expression data from peripheral blood is taken immediately before the challenge and at 12, 24, and 48 hours post-challenge.

Apr 2021 - Jul 2021 **Data Science Intern**, *Viettel Digital*, digital.viettel.vn.
- Completed courses on deep learning, computer vision, natural language processing, and big data.
- Participated in a Kaggle contest, achieved a ranking of 108/4410 on the final private leaderboard.
  - Predicted customers' loyalty score for payment brands in Brazil [contest, solution].

Publications

2023 **Continuous Neural Ordinary Differential Equations**.
Huyen K. Vo*, Nghia H. Nguyen*, Long M. Bui*, Tan M. Nguyen, Stanley J. Osher, Thieu N. Vo (In preparation).

2022 **Improving Neural Ordinary Differential Equations with Nesterov's Accelerated Gradient Method**.
Nghia H. Nguyen*, Tan M. Nguyen*, **Huyen K. Vo**, Stanley J. Osher, Thieu N. Vo (**NeurIPS**: Conference on Neural Information Processing Systems, 2022, **rank A***) [pdf].

2022 **Development of Vietnamese Text-To-Speech for VLSP Challenge 2021**.
Le Minh Nguyen, Do Quoc An, Vu Quoc Viet, **Vo Thuc Khanh Huyen** (**VNU Journal of Science**: Computer Science and Communication Engineering, 2022)[pdf].

National/International Prize

2019 **Vietnam Mathematics Olympiad (VMO)**, *Third prize (2019).*

2017 **Singapore Mathematics Olympiad (SMO)**, *Silver medal (2017).*

Achievements

2023 **Second Place in FPT Software AI Center Hackathon**, *Building LLM applications.*
   - Created a travel planner app using the ChatGPT API, Google reviews API and Langchain agents to design personalized itineraries based on customers' preferences.

2022 **Second Place in UET Hackathon**, *Track: Data Science, [contest, solution].*
   - Predicted the income based on people's demographic and work experience.

2021 **First Place in the shared task of Text To Speech, VLSP 2021**, *The 8th International Workshop on Vietnamese Language and Speech Processing, [contest, solution].*
   - Developed a novel Vietnamese text-to-speech model, achieving 3.729 in-domain MOS, 3.557 out-of-domain MOS and 79.70% SUS score.
   - Proposed an effective method to preprocess a dataset that only consists of spontaneous audio, trained a FastSpeech 2 acoustic model with some replacements in the external aligner model and used HiFiGAN vocoder to construct the waveform.

Other Experience

Oct 2022 - Present **Machine Learning Teaching Assistant**, MINDX TECHNOLOGY SCHOOL, mindx.edu.vn.
   - Helped develop curriculum for over 50 students.
   - Supported coding, technical issues, and graded student assignments.
   - Courses: machine learning (foundational algorithms, machine learning lifecycle involves data collection, preprocessing, feature engineering, model building, evaluation, and deployment).

Oct 2019 - Mar 2020 **Mathematics Teaching Assistant**, HEXAGON OF MATHS AND SCIENCE, hexagon.edu.vn.
   - Taught mathematics for students who attend Singapore University Entrance Exam (UEE).

Projects

2023 **Big Data Analysis of Binance market through Trade data and Twitter sentiment**.
   - A Big Data pipeline to aggregate Twitter data, perform users' sentiment analysis on different cryptocurrencies and correlate findings with their market prices.
   - Implemented Lambda architecture: (1) used Binance and Twitter stream API for data extraction, preprocessed, and pushed data to Kafka topics; (2) utilized Spark for processing batch data from HDFS and Spark Streaming for real-time updates; (3) stored results in Cassandra; (4) developed a visualization platform using Flask (back-end) and ReactJS (front-end).
   - Analyzed users' sentiment analysis using NLP techniques.

Technical Skills

Languages	Python, R, Java, C/C++, SQL.
Frameworks	PyTorch, Tensorflow, Keras, Scikit-learn.
Big Data	AWS, Spark, Hadoop, Kafka, NoSQL.
DevOps	Linux, Docker.
Web	HTML5, CSS3, JavaScript, jQuery, Ajax, Java (Servlet, Spring MVC, Spring Boot).
Certificates	Machine Learning DevOps Engineer Nanodegree program (Udacity), Data Engineering Nanodegree program (Udacity), Data Streaming (Udacity).

Languages

Vietnamese	Native.
English	IELTS Overall 7.5: Reading 8.5, Listening 8.0, Writing 6.5, Speaking 6.5.
Others	Japanese, Korean.

Extracurricular Activities

2023 **Vietnam Summer School of Science (VSSS) '10**, *Participant.*

2022 **Pre-PhD Summer School by VinUni-Illinois Smart Health Center**, *Participant.*

2021 - 2022 **Student Association's Educational Support Department, HUST**, *Member.*

2019 - 2021 **FPT Center for Young Talents (FYT)**, *Member.*

2020 **Hanoi Math Modelling**, *Content Team.*