Hoang Phan

646-508-0988 | phanviethoang1512@gmail.com | viethoang1512.github.io | hoang-pv | VietHoang1512

Education

New York University — New York, USA
Ph.D., Data Science, Advisors: Prof. Qi Lei, Prof. Andrew Gordon Wilson. — *Sept. 2023 - Present*

Hanoi University of Science and Technology — Hanoi, Vietnam
B.Sc., Computer Science, Honor program. — *Aug. 2018 - Aug. 2022*

Experience

VinAI Research — Hanoi, Vietnam
Research resident — *Aug. 2021 - Feb. 2023*
Advisor: Prof. Trung Le

- Main research topics: *Robust machine learning, Transfer learning, Multi-task learning, Self-supervised learning.*
- Participated in a Smart City project that aims to build a kidnapping and unshielded truck detection system.

Data Science Laboratory — SoICT-HUST
Undergraduate research student — *Sep. 2019 - Aug. 2021*
Advisors: Prof. Khoat Than, Prof. Huong Le Thanh, Dr. Linh Ngo Van

- Main research topics: *Probabilistic inference, Continual learning.*
- Developed a processing toolkit for the Khmer language (> 40.000 downloads and installations), supporting Asian Language machine translation. Used in Meta Research projects (NLLB, LASER, stopes), NAACL Workshop ...

Cinnamon Inc — Hanoi, Vietnam
AI research engineer — *Jan. 2020 - June. 2021*

- Worked on client productions, meta-learning and NLP-related research topics. Mainly focusing on information extraction of domain-specific business documents with the limitation of provided training data.
- Developed an internal usage chatbot, integrated with Slack, in support of the Human Resource department.

Publications

- **Hoang Phan**, Andrew Gordon Wilson, Qi Lei. "Controllable Prompt Tuning for Balancing Group Distributional Robustness". *International Conference on Machine Learning*, 2024. [pdf]
- Anh Nguyen, Long Vuong, **Hoang Phan**, Toan Do, Dinh Phung and Trung Le. "Flat Seeking Bayesian Neural Networks". *Advances in Neural Information Processing Systems*, 2023. [pdf]
- **Hoang Phan**, Trung Le, Trung Phung, Anh Bui, Nhat Ho and Dinh Phung. "Global-Local Regularization Via Distributional Robustness". In *International Conference on Artificial Intelligence and Statistics*, 2023. [pdf]
- **Hoang Phan**, Ngoc Tran, Trung Le, Toan Tran, Nhat Ho and Dinh Phung. "Stochastic Multiple Target Sampling Gradient Descent". In *Advances in Neural Information Processing Systems*, 2022. [pdf]
- **Hoang Phan***, Anh Phan*, Son Nguyen*, Linh Ngo Van and Khoat Than. "Reducing Catastrophic Forgetting in Neural Networks via Gaussian Mixture Approximation". In *Advances in Knowledge Discovery and Data Mining. PAKDD*. Lecture Notes in Computer Science. Springer, 2022. [pdf]
- **Hoang Phan**, Long Nguyen, Long Nguyen, and Khanh Doan. "Matching The Statements: A Simple and Accurate Model for Key Point Analysis". In *Proceedings of the 8th Workshop on Argument Mining* at EMNLP, 2021. [pdf]
- Tam Nguyen, Quang Pham, Linh Doan, Hoang Trinh, Anh Nguyen and **Hoang Phan**. "Contrastive Learning for Natural Language-Based Vehicle Retrieval". In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2021. [pdf]

Submissions

- Yijun Dong*, **Hoang Phan***, Xiang Pan* and Qi Lei. "Sketchy Moment Matching: Toward Fast and Provable Data Selection for Finetuning". Under review.
- **Hoang Phan***, Lam Tran*, Quyen Tran and Trung Le. "Enhancing Domain Adaptation through Prompt Gradient Alignment". Under review.
- **Hoang Phan**, Lam Tran, Ngoc Tran, Nhat Ho, Dinh Phung and Trung Le. "Improving Multi-task Learning via Seeking Task-based Flat Regions". Under review. [pdf]
- Ngoc Tran, Son Duong, **Hoang Phan**, Tung Pham, Dinh Phung and Trung Le. "Sharpness & Shift-Aware Self-Supervised Learning". Under review. [pdf]
- Ngoc Tran, Lam Tran, **Hoang Phan**, Anh Bui, Tung Pham, Toan Tran, Dinh Phung and Trung Le. "Robust Contrastive Learning With Theory Guarantee". Under review. [pdf]
- Quyen Tran, **Hoang Phan**, Khoat Than, Dinh Phung and Trung Le. "Continual Learning with Optimal Transport based Mixture Model". Under review. [pdf]
- Quang Pham, **Hoang Phan**, Hoang Trinh, Anh Nguyen and Hoai Nguyen. "S-NLP at ICDAR 2021: Multimodal Emotion Recognition on Comics scenes". Submitted to *International Conference on Document Analysis and Recognition Workshops*. [pdf]

(∗) denotes equal contribution

National/international prize

National Olympiad for Students 2017 - 2018
Vietnam Mathematical Society (VMS) Olympiad for universities and high school students
- Gold medal (2018), Silver medal (2017).

Vietnam Mathematics Olympiad (VMO) 2017 - 2018
The annual national mathematics competition for high school students
- Second prize (2018), Third prize (2017).
- International Mathematics Olympiad team selection test contestant.

International Tournament of Towns (ITOT) 2017
International mathematical Olympiad for school students
- Bronze medal (2017).

Vietnam Institute for Advanced Study in Mathematics (VIASM) Scholarship 2016 - 2018
Scholarship of the National Program for the Development of Mathematics

Vallet Scholarship 2017 - 2018
Prof. Odon Vallet's scholarships for outstanding Vietnamese students

Achievements

ArgMining: Key Point Analysis Shared Task, EMNLP 2021
Presented the 4^{th} solution in the workshop, hosted by IBM Research AI team.

NVIDIA AI City Challenge: Natural Language-Based Vehicle Retrieval, CVPR 2021
Got the 4^{th} place with a Siamese model capable of extracting visual concepts from natural language.

Multimodal Emotion Recognition on Comics scenes, ICDAR 2021
Won the first prize with a late fusion model, combined the processing of multiple input modalities.

Reliable Intelligence Identification on Vietnamese SNSs, VLSP 2020
Achieved a competitive performance (ranked 4^{th}) in the VLSP shared task.

Credit Scoring Challenge, Kalapa. Jsc 2020
Team ranked 1/146 with 800 contestants in constructing a banking credit scoring system.

University of Liverpool - Ion Switching, Kaggle 2020
Bronze medalist (solo) in the University of Liverpool's Institute of Ageing and Chronic Disease competition.

Practical deep learning, Vietnam Institute for Advanced Study in Mathematics 2019
Developed a deep learning-based Vietnamese spell corrector and ranked 1/50 in the final-project defense.

Professional services

Reviewer at NeurIPS, ICLR, ICCV, ECCV.