THANH NGOC (NATALIE) PHAM

874 Flemington St., Pittsburgh, PA, 15217

thanhngp@andrew.cmu.edu | (412) 785-4349 | Github | Website | LinkedIn

EDUCATION

Carnegie Mellon University (CMU), Pittsburgh, PA — *Aug 2022 - May 2023*

Master of Science, Data Analytics for Science, GPA: 4.12/4.33

Chinese University of Hong Kong (CUHK), Hong Kong — *Sep 2018 – Jul 2022*

Bachelor of Science, Mathematics, GPA: 3.05/4.00

University of Pittsburgh, Pittsburgh, PA — *Jan 2022 – Apr 2022*

Undergraduate semester exchange program, GPA: 4.00/4.00

RESEARCH EXPERIENCE

Machine Learning Department, CMU — Pittsburgh, PA

Research Assistant, advised by **Prof. Aarti Singh, Prof. Shirley Ho** and **Prof. Barnabas Poczos** — *Jan 2024 – Now*

- Experience recent deep learning structure like Mamba and Denoising Diffusion Probabilistic Model on learning partial differential equations' numerical solution and IllustrisTNG cosmological data.
- Improve existing Mamba structures for autoregressive training on partial differential equations' temporal inference with physics-informed loss.

Research Assistant, advised by **Prof. Andrej Risteski** — *April 2023 – Now*

- Investigated the S4 model and its variants for processing long-sequences and time series data.
- Analyze the performance of S4 and its variants, like Sashimi, and derive provable sub-linear regret bound for linear dynamical system in online learning setting with continuous signals.

Mathematics Department, CUHK — Hong Kong

Undergraduate Researcher, advised by **Prof. Eric Chung** — *Jan 2022 – Jul 2022*

- Studied data-driven reduced-order modeling for time-dependent problems.
- Used reduced-basis methods and data-driven approaches (such as Gaussian Process or Neural Networks) to approximate numerical solution to partial differential equations such as viscous Burger's equations.

Shanghai Jiao Tong University — Hong Kong - Virtual

Summer Research Intern, advised by **Prof. Lizhuang Ma** — *Jul 2021 – Aug 2021*

- Investigated 3D object detection algorithms based on point cloud and semi-supervised learning, and analyzed their advantages and disadvantages.
- Combined point-based neural networks such as PV-RCNN with a semi-supervised learning approach to resolve the issue of limited labeled data and validated the proposed model with KITTI dataset.

SELECTED PROJECTS

Streamline Protein Image Processing, CMU — *Jan 2023 – May 2023*

- Collaborated with a team of data scientists from Bristol Myers Squibb as part of master's capstone project.
- Generated pseudo-labels for 4000+ unlabeled protein crystallization images and HS-AFM data.
- Developed computational strategies using deep learning methods (rVAE, CNNs) to automate image processing and analysis workflows on multiple datasets.

Analyze Data from Public Reports of Asian giant hornets in Washington, CUHK — *Jan 2021 - Feb 2021*

- Constructed a model for predicting the existence of Asian giant hornets using public reports and reasonably allocate workforce to remove the hornets' nest for the Mathematical Contest in Modeling (MCM) 2021
- Applied the Metropolis Hastings algorithm to predict the spread of the hornets based on 14 positive public reports to update possible existence in a spatial distribution.
- Trained an image classifier with convolutional neural networks on Google Colab and achieved 85% accuracy on evaluating the likelihood of false reports submitted by the public.

TEACHING EXPERIENCE

Machine Learning Department, CMU — Pittsburgh, PA

Graduate Teaching Assistant — *Sep 2023 – Now*

- Assist the lecturer of the course 10-708: Probabilistic Graphical Models for designing assignments and grading.
- Hold recitations and office hours to clarify students' understanding of the course materials, including topics such as Markov Chain Monte Carlo, Variational Autoencoders, Causal Inference.

AWARDS

Yasumoto International Exchange Scholarship, **CUHK** — *Jan 2022*

Honours At Entrance, **CUHK** — *Sep 2018*

Faculty Admission Scholarship for Science Students, ***Science Faculty,*** **CUHK** — *Sep 2018*