Dang Nguyen

dangnguyen@uchicago.edu

Education

The University of Chicago   Chicago, IL
B.S. in Computer Science, Machine Learning specialization   Expected 2024
B.A. in Mathematics   Expected 2024

Research Interests

I am interested in vision-language representation learning, the intersection of causality and neural network explanations, and healthcare applications of deep learning.

Awards and Honors

UChicago Quad Research Grant   2023
UChicago Metcalf Internship Grant   Spring 2023, Fall 2023
UChicago Summer Match Internship Grant   2021
UChicago Dean's List   2021

Publications

Nguyen, D., Chen, C., He, H., & Tan, C. (2023, December). Pragmatic Radiology Report Generation. In Machine Learning for Health (ML4H) (pp. 385-402). PMLR.

Research Experience

Chicago Human+AI Lab, Computer Science Department   UChicago
Student Researcher   Jan 2023-Present
- Advisor: Dr. Chenhao Tan, Assistant Professor of Computer Science.
- Built a multi-modal model for effective chest X-ray report generation.
- Showed, using a Chi-squared test, that mentions of negative findings are driven by the X-ray study indication.
- Introduced a method to prompt Flan-T5 to clean ~200,000 reports, reducing model hallucination by 69.5% from the previous state-of-the-art.
- Developed a vision-language neural network that outperforms the previous SoTA by 29% on Positive F1, 43% in BERTScore, and 52.9% on Negative F1.

UCARE Lab, Computer Science Department   UChicago

Student Researcher   Dec 2020-Sep 2022

- Advisor: Dr. Haryadi Gunawi, Associate Professor of Computer Science.
- The project aims to develop a machine learning model for slow disk detection.
- Built a pipeline to retrieve and parse disk performance data from industry partner's database, and optimized it by 99%.
- Applied K-means and DBSCAN on 23 out of 91 pairwise correlations of data features to detect anomalous disk clusters.
- Developed an algorithm that uses KL divergence to detect slow disks and analyzed ~50,000 disks in 3 days using 3 high-performance compute servers.

Work Experience

NetApp, Inc.   Remote

Research Intern   Jun-Sep 2022

- Worked on the Slow Disk Detection project to detect slow HDDs using machine learning.
- Optimized the custom algorithm based on KL divergence and improved NetApp's threshold-based detector's true positive rate from 10% to 59%.
- Sped up the model evaluation process by compiling 3 data sets of slow disks using Apache Hive.

Teaching/Mentorship

Maroon Tutor Match   Chicago, IL

Math/General Subjects Tutor   Mar 2021-May 2022

- Tutored three middle schoolers one-on-one, three hours per student per week.
- Improved a 7th-grader's ISEE test's Quantitative Reasoning percentile from 2nd to 47th, and Mathematics Achievement from 7th to 27th.
- Taught algebra to 6th, 7th graders to assist them with homework, and helped an 11th grader with SAT preparation.
- Taught writing to 6th, 7th graders and provided feedback on their argumentative essays.

Presentations

Machine Learning for Health   New Orleans, LA, Dec 2023

Poster title: *Pragmatic Radiology Report Generation*

Midstates Undergraduate Research Symposium   Chicago, IL, Nov 2023

Poster title: *Pragmatic Radiology Report Generation*

Symposium on Human+AI   Chicago, IL, Sep 2023

Poster title: *Pragmatic Radiology Report Generation*

NetApp Intern Final Presentation   Remote, Sep 2022

Poster title: *Slow Disk Detection*

Relevant Coursework

Machine Learning/Statistics: Machine Learning, Fundamentals of Deep Learning (graduate-level), Statistical Models and Methods, Introduction to Mathematical Probability.
Mathematics: Real Analysis I-III, Abstract Linear Algebra, Abstract Algebra I-II, Discrete Mathematics, Complexity Theory (in-progress).
Computer Science: Theory of Algorithms, Functional Programming, Computational Geometry, Introduction to Database Systems, Computer Architecture (in-progress).

Languages

Natural: fluent in English, native in Vietnamese, intermediate in Mandarin.
Programming: advanced in Python (especially statistical, ML, and DL libraries), intermediate in C and Java, beginner in JavaScript.

Technical Skills

Data Science/ML/DL: Pandas, Scikit-learn, NumPy, Matplotlib, PyTorch, Huggingface.
Front-end web development: HTML/CSS, ReactJS.
Back-end web development: NodeJS, ExpressJS, MongoDB, Firebase, SQL.

References

Chenhao Tan
Assistant Professor of Computer Science
University of Chicago
chenhao@uchicago.edu

Haryadi Gunawi
Associate Professor of Computer Science
University of Chicago
haryadi@cs.uchicago.edu

He He
Assistant Professor of Computer Science
New York University
hhe@nyu.edu