Vincent Foster Chen

vfchen2@illinois.edu | +1(310)994-3273 | linkedin.com/in/vincent-chen2/

Education

University of Illinois at Urbana-Champaign *Completed Dec 2023*
Master of Computer Science *GPA: 4.0/4.0*
Relevant Courses : Distributed Systems, Database Systems, Applied Parallel Programming

National Tsing Hua University (NTHU) *Completed Jun 2021*
Bachelor of Science in Physics, Minor in Computer Science *GPA: 3.74/4.0*
Relevant Courses: Algorithms, Data Structures, Operating Systems, Computer Architecture

Technical Skills

Programming Languages: C/C++, Python, R, MySQL, CUDA, Javascript, Shell Script
Packages & Tools: PyTorch, Keras, Scikit-learn, Tensorflow, Linux, Hadoop, Git, MongoDB, Docker, MLflow
Certifications: Artificial Intelligence Program (NTHU), Data Science Program (NTHU)

Work Experience

Robotics and Electromechanical Systems, BlueHalo *May 23 – Aug 23*
Machine Learning Intern *Rockville, Maryland*

- Developed a Transformer-based model using MLflow to render complex scenes for few-shot view synthesis
- Performed data collection with MongoDB and PyTorch to gather a large scale dataset with 8 million images
- Streamlined data collection pipelines within Docker containers, enhancing data transfer across multiple servers

Natural Language Processing and Sentiment Analysis Lab, Academia Sinica *May 21 – May 22*
Research Assistant *Taipei, Taiwan*

- Published paper, "Learning to Rank Visual Stories with Human Ranking Data" in the Association for Computation Linguistics (ACL 2022)
- Published paper, "Multi-VQG: Generating Engaging Questions for Multiple Images" in the Conference on Empirical Methods in Natural Language Processing (EMNLP 2022)
- Created a novel and generalizable SOTA metric for visual storytelling with PyTorch
- Designed UI with HTML, CSS, and JavaScript to collect multi-image question dataset with Amazon Mechanical Turk

Acoustic Hearing Group, NTHU *Feb 20 – Aug 21*
Undergraduate Research Assistant *Hsinchu, Taiwan*

- Utilized the Kaldi tool kit with Shell Script to perform adaptation on Automatic Speech Recognition
- Implemented Transformer Model from scratch with Tensorflow to create a text machine translator
- Debugged pre-designed Python algorithms of video subtitle generator to extend video recognition duration

Selected Projects

Hallucination Analysis with Small Language Models *Sep 23 – Dec 23*

- Created a synthetic dataset which allows small language models(LM) to learn simple facts
- Proposed a benchmark to probe LMs for factual hallucinations and produce hallucination rates
- Performed analysis of hallucination emergence by comparing LMs between scale and hallucination rate

MapleJuice Cloud Computing *Sep 23 – Dec 23*

- Built a cloud computing framework which performs 5% faster than Hadoop on small cluster
- Developed a distributed file system supporting get, put, delete operations with replica control
- Utilized gossiping failure detection to maintain a dynamic P2P membership list with fault tolerance

Parameter Efficient Fine-Tuning *Feb 23 – May 23*

- Conducted a comprehensive analysis of lightweight tuning methods applied to LM
- Optimized the loss landscapes of LMs, resulting in SOTA performance on various downstream tasks

Optimization of Convolutional Neural Network with CUDA *Sep 22 – Dec 22*

- Implemented neural-network convolutional layers forward pass with various optimization techniques
- Conducted analysis of CUDA kernels using Nsight-System and Nsight-Compute profiling tools

Deep Learning Research *Sep 21 – Nov 21*

- Published paper "Associated Learning: an Alternative to End-to-End Backpropagation that Works on CNN, RNN, and Transformer" in the International Conference of Learning Representation (ICLR 2022)
- Researched and implemented an alternative of back-propagation using Tensorflow