Mengxi Wu

mengxiwu@usc.edu
wmx567.github.io

Education

University of Southern California Expected 2027.12
Doctor of Philosophy, Computer Science

University of Southern California Expected 2027.5
Master of Arts, Pure Mathematics

New York University 2021.5
Master of Science, Computer Science

University of Michigan, Ann Arbor 2019.5
Bachelor of Science in Engineering, Electrical Engineering

Experience

USC Information Science Institute Present
PhD Student, Advised by Prof.Xuezhe Ma

- Conducted research on efficient pre-training of large language models (LLMs) and understanding the mathematical principles underlying training dynamics, including the hyperparameter optimization and optimizers.

Huawei 2021.9 - 2022.5
Machine Learning Engineer, Network AI Engine Department

- Developed intelligent systems to detect fire, smoke, and helmet in 2D images captured from cameras on construction sites.
- Integrated detection models with techniques such as Knowledge Distillation and Memory Replay to enable class-incremental learning.

NYU Multimedia and Visual Computing Lab 2021.3 - 2022.5
Research Assistant, Advised by Prof.Yi Fang

- Conducted research on geometric-aware adversarial training methods for 3D point cloud completion.
- Designed a novel adversarial attack method that constrains adversarial perturbations with absolute minimum curvature direction of original data.

NYU Algorithms and Foundations Group 2020.6–2022.4
Research Assistant, Advised by Prof.Yi-Jen Chiang and Prof.Christopher Musco

- Conducted research on in situ selection of key time steps for high dimensional time-varying data.
- Developed a new greedy algorithm with numerical linear algebra techniques to compute linear interpolation solutions and errors in an online streaming fashion.

Publications

(∗ indicates equal contribution.)

7. **Towards Self-Adaptive Learning: Continual Learning under Harsh Conditions.**
   Ehsan Hallaji, Alireza Fathalizadeh, Mengxi Wu, Mohammad Rostami, Roozbeh Razavi-Far.
   Submitted
6. **Gecko: An Efficient Neural Architecture Inherently Processing Sequences with Arbitrary Lengths.**
   Xuezhe Ma, Shicheng Wen*, Linghao Jin*, Bilge Acun*, Ruihang Lai*, Bohan Hou, Will Lin, Hao Zhang, Songlin Yang, Ryan Lee, Mengxi Wu, Jonathan May,*

Luke Zettlemoyer, Carole-Jean Wu.
Preprint

5. **GQA-μP: The Maximal Parameterization Update for Grouped Query Attention and Fully Sharded Data Parallel.**
Kyle R. Chickering, Huijuan Wang*, Mengxi Wu*, Alexander Moreno, Muhao Chen, Xuezhe Ma, Daria Soboleva, Joel Hestness, Zhengzhong Liu, Eric P. Xing.*
Submitted

4. **Curvature Diversity-Driven Deformation and Domain Alignment for Point Cloud.**
Mengxi Wu, Hao Huang, Yi Fang, Mohammad Rostami.
Transactions on Machine Learning Research 2025

3. **Graph Harmony: Denoising and Nuclear-Norm Wasserstein Adaptation for Enhanced Domain Transfer in Graph-Structured Data.**
Mengxi Wu, Mohammad Rostami.
Transactions on Machine Learning Research 2024

2. **Streaming Approach to In Situ Selection of Key Time Steps for Time-Varying Volume Data.**
Mengxi Wu, Yi-Jen Chiang, Christopher Musco.
Eurographics/IEEE Conference on Visualization 2022

1. **3D Point Cloud Completion with Geometric-Aware Adversarial Augmentation.**
Mengxi Wu, Hao Huang, Yi Fang.
International Conference on Pattern Recognition 2022

Teaching

University of Southern California
CSCI 544 Applied Natural Language Processing, *Teaching Assistant*
CSCI 570 Analysis of Algorithms, *Teaching Assistant*

New York University
ECE-GY 9123 Deep Learning, *Teaching Assistant*

Services

Conference Reviewing
EMNLP, ICLR, ICML

Journal Reviewing
Transactions on Machine Learning Research

Skills

Programming Languages
C/C++, Python, Java, R, MATLAB, Swift

Libraries and Tools
PyTorch, Tensorflow, PySpark, Hadoop