Zheda Mai

: mai.145@osu.edu : 614-967-1380 : zheda-mai.github.io

EDUCATION

Ph.D. Computer Science and Engineering, Ohio State University 2022-2027 (Expected)
- Research areas: ***Efficient Foundation Model Adaptation***, ***Multimodal LLM***, ***Continual Learning***, ***Learning with Imperfect Data***
- Advisor: Professor Wei-Lun Chao. GPA: 3.96/4.0
- Google Scholar: 1100+

M.A.Sc. Information Engineering, University of Toronto 2018-2021
- Research areas: Continual Learning, Recommender Systems
- Advisor: Professor Scott Sanner. GPA: 4.0/4.0

B.A.Sc. Engineering Science, University of Toronto 2012-2017

PUBLICATIONS

* denotes equal contributions and co-first authorship.

Conferences

[NAACL 2025] Zhongwei Wan, Hui Shen, Xin Wang, Che Liu, **Zheda Mai**, Mi Zhang. Attention Entropy-Guided Dynamic Cache Allocation for Efficient Multimodal Long-Context Inference. In *Proceedings of Annual Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics (NAACL)*, 2025.

[ICASSP 2025] Jiageng Zhu, Kehao Li, **Zheda Mai**, Hanchen Xie, Wael AbdAlmageed, Zubin Abraham. Attention-Driven Causal Discovery: From Transformer Matrices to Granger Causal Graphs for Non-Stationary Time-series Data. In *Proceedings of International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2025.

[NeurIPS 2024] **Zheda Mai***, Jihyung Kil*, Justin Lee, Zihe Wang, Kerrie Cheng, Lemeng Wang, Ye Liu, Arpita Chowdhury, Wei-Lun Chao. MLLM-CompBench: A Comparative Reasoning Benchmark for Multi-modal LLMs. In *Proceedings of the Conference on Neural Information Processing Systems (NeurIPS)*, 2024.

[NeurIPS 2024] **Zheda Mai***, Arpita Chowdhury*, Ping Zhang*, Cheng-Hao Tu, Hong-You Chen, Vardaan Pahuja, Tanya Berger-Wolf, Song Gao, Charles Steward, Yu Su, Wei-Lun Chao. Fine-Tuning is Fine, if Calibrated. In *Proceedings of the Conference on Neural Information Processing Systems (NeurIPS)*, 2024.

[NeurIPS 2023] Cheng-Hao Tu*, Hong-You Chen*, **Zheda Mai**, Jike Zhong, Vardaan Pahuja, Tanya Berger-Wolf, Song Gao, Charles Steward, Yu Su, Wei-Lun Chao. Holistic Transfer: Towards Non-Disruptive Fine-Tuning with Partial Target Data. In *Proceedings of the Conference on Neural Information Processing Systems (NeurIPS)*, 2023.

[CVPR 2023] **Zheda Mai***, Cheng-Hao Tu*, Wei-Lun Chao. Visual Query Tuning: Towards Effective Usage of Intermediate Representations for Parameter and Memory Efficient Transfer Learning.In *Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023.

[WWW 2022] **Zheda Mai***, Tianshu Shen*, Ga Wu, Scott Sanner. Distributional Contrastive Embedding for Clarification-based Conversational Critiquing. In *Proceedings of the ACM Web Conference (WWW)*, 2022.

[SIGIR 2022] Zhaolin Gao, Tianshu Shen, **Zheda Mai**, Mohamed Reda Bouadjenek, Scott Sanner. Mitigating the Filter Bubble while Maintaining Relevance: Targeted Diversification with VAE-based Recommender Systems. In *Proceedings of Special Interest Group on Information Retrieval (SIGIR)*, 2022.

[AAAI 2021] **Zheda Mai***, Dongsub Shim*, Jihwan Jeong*, Scott Sanner, Hyunwoo Kim, Jongseong Jang. Online Class-Incremental Continual Learning with Adversarial Shapley Value. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, 2021.

Journals

[JVCI 2023] Ruiwen Li, **Zheda Mai**, Chiheb Trabelsi, Zhibo Zhang, Jongseong Jang, Scott Sanner. TransCAM: Transformer Attention-based CAM Refinement for Weakly Supervised Semantic Segmentation. In *Journal of Visual Communication and Image Representation (JVCI)*, 2023.

[IPM 2023] Tianshu Shen, Jiaru Li, Mohamed Reda Bouadjenek, **Zheda Mai**, Scott Sanner. Unintended Bias in Language Model-driven Conversational Recommendation. In *Information Processing and Management (IPM)*, 2023.

[Neurocomputing 2022] **Zheda Mai**, Ruiwen Li, Jihwan Jeong, David Quispe, Hyunwoo Kim, Scott Sanner. Online Continual Learning in Image Classification: An Empirical Survey. In *Neurocomputing*, 2022.

[AIJ 2022] Vincenzo Lomonaco, ..., **Zheda Mai**, etc. CVPR 2020 Continual Learning in Computer Vision Competition: Approaches, Results, Current Challenges and Future Directions. In *Artificial Intelligence Journal (AIJ)*, 2022.

Workshops

[NeurIPS 2023] **Zheda Mai***, Tianle Chen*, Ruiwen Li, Wei-Lun Chao. Segment Anything Model (SAM) Enhanced Pseudo Labels for Weakly Supervised Semantic Segmentation. In *Proceedings of the Conference on Neural Information Processing Systems (NeurIPS) Workshops*, 2023.

[CVPR 2021] **Zheda Mai**, Ruiwen Li, Hyunwoo Kim, Scott Sanner. Supervised Contrastive Replay: Revisiting the Nearest Class Mean Classifier in Online Class-Incremental Continual Learning. In *Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2021.

[CVPR 2020] **Zheda Mai**, Hyunwoo Kim, Jihwan Jeong, Scott Sanner. Batch-level Experience Replay with Review for Continual Learning. In *Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2020.

[ICDM 2020] **Zheda Mai***, Ga Wu*, Kai Luo, Scott Sanner. Attentive Autoencoders for Multifaceted Preference Learning in One-class Collaborative Filtering. In *Proceedings of International Conference on Data Mining (ICDM) Workshops*, 2020.

Under Review

[R1] **Zheda Mai**, Ping Zhang, Cheng-Hao Tu, Hong-You Chen, Li Zhang, Wei-Lun Chao. Lessons Learned from a Unifying Empirical Study of Parameter-Efficient Transfer Learning (PETL) in Visual Recognition, 2024

Technical Reports

[T1] JinPeng Zhou, Ga Wu, **Zheda Mai**, Scott Sanner. Noise Contrastive Estimation for Autoencoding-based Collaborative Filtering, 2020.

EXPERIENCE

Research Intern, *Bosch Research*, Sunnyvale — May 2024 - Aug. 2024
- Developed a unified multimodal framework for general time series analysis using language and vision foundation models.

Data Scientist, *Optimy AI*, Canada — 2021-2022
- Developed machine learning models for customer engagement and purchase likelihood predictions.

Machine Learning Engineer Intern, *Pitney Bowes*, Canada — May 2019 - Oct. 2019
- Developed map-style extraction models with CNN and multi-task learning.

Software Engineer Intern, *AMD*, Canada — 2015 - 2016
- Developed design verification tool automation for Verilog.

AWARDS

- Outstanding Reviewer Award for **NeurIPS** 2023 — 2023
- **4th place** of the CLVision Continual Learning challenge at **CVPR** 2021 — 2021
- **1st place** of the CLVision Continual Learning challenge at **CVPR** 2020 — 2020

SKILLS

Techniques: Python, SQL, Git, LaTex, AWS, PySpark, JavaScript
Machine Learning Tools: PyTorch, Keras, TensorFlow, NumPy, Pandas, SciPy, scikit-learn

PROFESSIONAL SERVICE

I am a conference reviewer for

- NeurIPS: 2023, 2024
- ICML: 2023, 2024
- ICLR: 2024, 2025
- CVPR: 2024, 2025
- IJCAI: 2024

I am a journal reviewer for

- ACM Computing Surveys
- Artificial Intelligence (AIJ)
- Frontiers in Artificial Intelligence

TALKS

- Continual Learning in Image Classification. Vector Institute. July 2020
- Recent Advances in Continual Learning. D3M Lab, University of Toronto Jan. 2022

TEACHING

Teaching Assistant, University of Toronto

- APS1070: Foundations of Data Analytics and Machine Learning (2019, 2020)
- MIE451: Decision Support Systems (2019, 2020)
- MIE1628: Big Data Science (2020)

VOLUNTEER

Judge and mentor

- Artificial Intelligence hackathon hosted by Artificial Intelligence Club of OSU: 2022, 2023, 2024