

Yuchen Shen

yuchens3@cs.cmu.edu | [website](#) | [Google Scholar](#)

Education

Carnegie Mellon University

2023/08-2024/12 (Expected)

Master of Science in Intelligent Information Systems, Language Technologies Institute

Pittsburgh, PA, USA

Courses: Probabilistic Graphical Models, Multimodal Machine Learning, Advanced Deep Learning

University of Electronic Science and Technology of China

2019/08-2023/06

Bachelor of Engineering in Software Engineering, School of Information and Software Engineering

Chengdu, China

Courses: Probability and Mathematical Statistics, Calculus I & II, Digital Signal Processing, Discrete Mathematics

Publications

- [1] **Yuchen Shen**, Haomin Wen, Leman Akoglu. Zero-shot Outlier Detection via Prior-data Fitted Networks: Model Selection Bygone!. Arxiv:2409.05672. (*under review at ICLR 2025*)
- [2] **Yuchen Shen**^{*}, Chenhao Zhang^{*}, Sijie Fu^{*}, Chenghui Zhou, Newell Washburn, Barnabás Póczos. Chemistry-Inspired Diffusion with Non-Differentiable Guidance. Arxiv:2410.06502. (*under review at ICLR 2025, accepted by ICML 2024 AI for Science Workshop.*)
- [3] **Yuchen Shen**, Barnabás Póczos. GraphBPE: Molecular Graphs Meet Byte-Pair Encoding. In ICML 2024 AI4Science Workshop.
- [4] Zhichao Hou, Weizhi Gao, **Yuchen Shen**, Feiyi Wang, Xiaorui Liu. Protransformer: Robustify Transformers via Plug-and-Play Paradigm. NeurIPS 2024. (*accepted* by ICLR 2024 Workshop on Reliable and Responsible Foundation Models).
- [5] **Yuchen Shen**, Xiaojun Wan. OpinSummEval: Revisiting Automated Evaluation for Opinion Summarization. Arxiv:2310.18122.
- [6] Fei Zhao^{*}, **Yuchen Shen**^{*}, Zhen Wu, Xinyu Dai. Label-Driven Denoising Framework for Multi-Label Few-Shot Aspect Category Detection. Findings of the Association for Computational Linguistics: EMNLP 2022.
- [7] Mengjuan Liu, Xiaoming Bao, Jiang Liu, Pei Zhao, **Yuchen Shen**. Generating emotional response by conditional variational auto-encoder in open-domain dialogue system. Neurocomputing, 460:106–116, 2021.

Research Experience

Task-Optimized (Visual-)Tactile Models with Feedback, Carnegie Mellon University

2024/09 – present

Advisor: Aran Nayebi, Assistant Professor, Machine Learning Department, School of Computer Science.

- Proposed to investigate how feedback improves classification accuracy for (visual-)tactile neural networks.
- Proposed to develop a feedback-enhanced model as the computational tool for the somatosensory cortex by matching the brain data from mice, and to identify the inductive bias needed for building next-generation (visual-)tactile neural networks.
- Developed PyTorchTNN, a Python package that unrolls neural networks over time with feedback to mimic biological systems.

Graph Generative Pre-training, Carnegie Mellon University

2024/05 – present

Advisor: Leman Akoglu, Associate Professor, Heinz College of Information Systems and Public Policy.

- Proposed to pre-train models with the auto-regressive graph generation objective on molecules for better representations.
- Investigated how dataset sizes, model architectures, and pre-training methods affect the quality of pre-training on molecules.

Synthetic Pre-training for Zero-Shot Anomaly Detection, Carnegie Mellon University

2024/05 – 2024/09

Advisor: Leman Akoglu, Associate Professor, Heinz College of Information Systems and Public Policy.

- Proposed to pre-train Prior-Data Fitted Network (PFN) on synthetic data and detect anomalies without further fine-tuning.
- Achieved a 20x speed-up for pre-training via data reuse with linear transformations on Gaussian mixture model priors.
- Our zero-shot model performs 2nd best out of 26 fine-tuned baselines on the Anomaly Detection Benchmark (ADBench) consisting of 57 datasets, with an inference speed of 7.7 ms/instance on average. *Paper under review at ICLR 2025.*

Geometry Optimized Molecule Generation with Chemistry Guidance, Carnegie Mellon University

2024/01 – 2024/05

Advisor: Barnabás Póczos, Associate Professor, Machine Learning Department, School of Computer Science.

- Proposed to optimize the stability of molecules with a latent diffusion model and GFN2-xTB, a semi-empirical quantum mechanical method for accurate and efficient quantum chemistry calculations, such as the forces on the atoms.
- Developed a method with zeroth-order optimization to estimate the gradient from a non-differentiable oracle (e.g., software) for guided diffusion on 3D molecules. Demonstrated the compatibility of the method in a bilevel molecular optimization framework.
- Achieved a decrease in force (up to 14.16%) and energy (up to 15.78%) with improved validity (up to 2.60%) for generated molecules on the QM9 and GEOM datasets. *Paper under review at ICLR 2025.*

Graph Tokenization for Molecules, Carnegie Mellon University

2024/01 – 2024/05

Advisor: Barnabás Póczos, Associate Professor, Machine Learning Department, School of Computer Science.

- Proposed to mine common subgraphs based on the topological structures and node types of molecules from a large corpus.

- Generalized Byte-Pair Encoding (BPE) algorithm to 2D graphs and developed a non-parametric tokenization method that partitions 2D graphs into subgraphs, which transforms the graph into a new 2D graph and a hypergraph at each tokenization step.
- Improved the performance for different GNNs and HGNNs (for hypergraphs) on both classification and regression datasets across various hyperparameters. Findings demonstrate the importance of data preprocessing in molecular machine learning.

Controllable Toxicity Generation for Plant Molecules, Carnegie Mellon University

2023/09 – 2023/12

Advisor: Barnabás Póczos, Associate Professor, Machine Learning Department, School of Computer Science.

- Constructed a dataset consisting of 1535 toxic and 39576 non-toxic molecules from plants.
- Achieved a classification accuracy of 80.8% on a balanced test set (1:1) for an imbalanced training set (~1:35). Proposed to achieve controllable toxicity generation via diffusion models with contrastive learning at both molecule-level and model-level.

Convergence of Decentralized Optimization Algorithms, North Carolina State University

2023/02 – 2023/12

Advisor: Xiaorui Liu, Assistant Professor, Department of Computer Science, College of Engineering.

- Analyzed convergence rates for different decentralized algorithms under convex and strongly convex settings.
- Studied and generalized the previous method to analyze lower-bound for decentralized optimization with (strongly) convex objectives. Improved the tightness of previous proofs on the convergence of decentralized algorithms such as [LEAD](#).

Automated Metric Evaluation for Opinion Summarization, Peking University

2022/12 – 2023/05

Advisor: Xiaojun Wan, Professor, Wangxuan Institute of Computer Technology.

- Proposed to evaluate automated metrics with aspect relevance, self-coherence, sentiment consistency, and readability.
- Constructed a dataset with annotated outputs from 14 popularly used models in opinion summarization.
- Analyzed 26 popularly used automatic metrics, showing neural-based metrics correlate better with human annotations.

Zero-Shot Unsupervised Opinion Summarization with Prefix-Tuning, Peking University

2022/02 – 2022/12

Advisor: Xiaojun Wan, Professor, Wangxuan Institute of Computer Technology.

- Proposed a zero-shot setting for unsupervised opinion summarization with unseen aspects to summarize in the test stage.
- Proposed to control the number of aspects and sentiment coherency in the generated summary with two kinds of prefixes.
- Improved the zero-shot performance on the [Space](#) dataset compared with strong baselines (e.g., on the “service” aspect, ROUGE-1 score improved from 33.56 to 35.94 compared with previous fine-tuned SOTA model [AceSum](#)).

Label-enhanced Few-shot Learning for Multi-label Aspect Category Detection, Nanjing University

2021/08 – 2022/01

Advisor: Xinyu Dai, Professor, Department of Computer Science and Technology.

- Identified the generic and noisy features as the bottleneck for multi-label few-shot aspect category detection (FS-ACD).
- Proposed to denoise the feature of each category with label texts. Designed a label-guided attention module and a label-weighted contrastive loss for FS-ACD to learn representative features and to distinguish semantically close categories.
- Improved the performance of existing models on FS-ACD (e.g., F1 score for the 5-way 5-shot setting, i.e., 5 classes with 5 instances to learn per class, improved from 75.37 to 78.27 on the [FewAsp](#) dataset for model [Proto-AWATT](#)).

Projects

Edge Weighting Algorithm with Ollivier-Ricci Curvature for Graph Classification

2023/09 – 2023/12

Carnegie Mellon University, 10-708 Probabilistic Graphical Models

- Proposed to weight edges based on Ollivier-Ricci curvature to overcome over-smoothing and over-squashing.
- Proposed to optimize the idleness in the computation of the Ollivier-Ricci curvature and achieved a classification accuracy of 70.90 ± 0.047 (with a baseline accuracy of 69.59 ± 0.048) on the Proteins dataset by weighting edges.

Multimodal Low-Rank Adaptation

2023/09 – 2023/12

Carnegie Mellon University, 11-777 Multimodal Machine Learning

- Proposed to share $n\%$ of the Low-Rank Adaptation (LoRA) matrices for texts and images in vision language models.
- Developed a regularization loss based on singular value decomposition (SVD) that promotes the similarity of LoRA matrices if more information is shared between the input language and vision modalities.
- Improved the performance of LoRA on 6 out of 7 subtasks of the VALSE benchmark (e.g., our method improved the pairwise rank accuracy of vanilla LoRA from 44.2% to 49.0% on the coreference-clean data for the CLIP model).

Honors

- Finalist Winner (Top 2%) in 2022 Mathematical Contest in Modeling (MCM) / Interdisciplinary Contest in Modeling (ICM)
- Honorable Mentions (Top 18%) in 2021 Mathematical Contest in Modeling (MCM) / Interdisciplinary Contest in Modeling (ICM)
- First Class Scholarship in University of Electronic Science and Technology of China (2020 2021, 2021-2022)

Skills

Language: Chinese (native), English (fluent)

Programming: Python (proficient), C, Java