Yongkang Du

Mobile: (+1)213-610-1669 Email: duyongka@gmail.com Website: https://yongkdu.github.io/

EDUCATION

Pennsylvania State University, **Aug 2024-Present**
Ph.D. Informatics
Advisor: Lu Lin

University of Southern California, **Jan 2022-May 2023**
M.S. Computer Science
Advisor: Jieyu Zhao

North China University of Technology, **Sep 2017-May 2021**
B.Eng. Computer Science and Technology, GPA 89.67/100, rank 4/154

PUBLICATION

- **FairCode: Evaluate Social Bias of LLM in Code Generation**
  ***Du, Y.**, Huang J., Zhao J., Lin L.*
  Preprint 2025
- **Mitigating Topology Bias in Graph Diffusion via Counterfactual Intervention**
  *Wang, W., **Du, Y.,** Yang, J., Lin L.*
  In submission KDD 2025
- **Self-Contradictory Reasoning Evaluation and Detection**
  *Liu, Z., Lee, I., **Du, Y.**, Sanyal S., Zhao, J.*
  EMNLP 2024 Findings
- **Controllable Pareto Trade-off Between Fairness and Accuracy**
  ***Du, Y.**, Zhao, J., Yang, Y., Zhou, T.*
  SoCal NLP Symposium 2023
- **FF-PRec: A Feature Fusion-Based Paper Recommendation Method on Academic Big Data of Heterogeneous Network**
  ***Du, Y.**, Ding, W., and Liang, Y.*
  2021 4th International Conference on Computer Information Science and Artificial Intelligence.

RESEARCH EXPERIENCE

Research Assistant, Penn State, **May 2024-Present**
Advisor**:** Lu Lin
Topic: Evaluate and Mitigate Social bias of LLM

- Propose a new benchmark to evaluate social bias in code generation
- Apply both real-world statistic data and synthetic code in diverse scenarios
- Quantify the bias issue with both refusal rate and preference entropy

Research Assistant, USC&UMD, **May 2023-Feb 2024**
Advisor**:** Jieyu Zhao**,** Tianyi Zhou
Topic: Multi-objective optimization, Fairness Accuracy Trade-off

- Purpose a novel gradient-based MOO algorithm with reference vector to get controllable trade-offs
- Utilize moving average of stochastic gradient for multiple-gradient descent algorithm, which gives a more precise common descent vector and lead to a better convergence
- Prune the high-dimensional noisy gradient with a mask generated based on parameter magnitude to reduce the negative impact of curse of dimensionality
- Compared with SOTA MOO methods, our method can best follow the reference vector and get the most controllable results on both training and test set

Research Assistant, Information Science Institute, USC, **Apr 2022-Aug 2022**
Advisor: Mohammad Rostami
Topic: Unsupervised Domain Adaptation for Event Camera

- Enhance the model's adaptation ability from image domain to event domain via adversarial training
- Apply Sliced Wasserstein Distance loss for generator to align the domain gap and generate better histogram representation for event data
- Utilize contrastive loss (InfoNCE) to pull the representations of samples of the same class across all domains closer while pushing representations of samples of different classes further
- Achieve 0.84 accuracy on N-Caltech dataset, which is 0.3 higher than the backbone model

Research Intern, Institute of Computing Technology, Chinese Academy of Sciences, Oct 2020-Nov 2021
Advisor: Ying Liang

Topic: Graph Representation Learning, Recommender System

- Purpose a new paper recommendation method in academic social network with deep learning algorithm
- Construct heterogeneous academic network that connects scholars, papers, and conferences
- Design a metapath-based graph representation learning algorithm to extract feature according to semantic information, which effectively alleviate the data sparsity problem
- Conduct experiments on the AMiner dataset, our method achieves 0.877 Hit Rate and 0.727 NDCG, which are 0.1-0.2 higher than the baseline methods

SERVICE & INVOLVEMENT

Volunteer Reviewer, CVPR 2023
Reviewer, EURASIP Journal on Wireless Communications and Networking
Reviewer, International Journal of Web Information Systems
Reviewer, Security and Communication Networks
Reviewer, EAI CollaborateCom 2024
Teaching Assistant, CSCI 544 Natural Language Processing, University of Southern California

HONERS & AWARDS

- Outstanding Undergraduate Thesis, North China University of Technology, 2021
- First-class Scholarship, North China University of Technology, 2018-2020
- First Prize in English Speech Contest, North China University of Technology, 2019

SKILLS

- Programming: Python, C/C++, Git, PyTorch, Numpy, TensorFlow, Pandas
- Language: Mandarin (native), English (advanced)