Yu Wang

Email: wy2001@mail.ustc.edu.cn | Google Scholar | Homepage

EDUCATION

University of Science and Technology of China *Aug. 2018 ∼ present*
School of the Gifted Young (Skipped the third year in high school)
Overall GPA: 3.95/4.3 (91.42/100) **Ranking**: 2/28
Major GPA: 4.11/4.3 (95.30/100)
Course Highlights: Data Structures(93), Introduction to Computer Systems(H)(96), Probability and Mathematical Statistics A(95), Fundamentals of Artificial Intelligence(99), Natural Language Processing(95), Computer Vision(95), Enterprise AI(98), Foundations of Algorithms(94), Operations Research(92)

RESEARCH INTERESTS

Information Retrieval; Machine Learning; Statistics; Reinforcement Learning; Causal Inference; Model Robustness; Meta-Learning; Outlier Detection; Domain Generalization.

PUBLICATIONS[1]

Meta-learning with an Adaptive Task Scheduler. PDF
Huaxiu Yao*, **Yu Wang***, Peilin Zhao, Mehrdad Mahdavi, Defu Lian, Ying Wei, Chelsea Finn
Accepted to NeurIPS 2021. (Accept rate: 26%)

Learning Robust Recommenders through Cross-Model Agreement. PDF
Yu Wang, Xin Xin, Zaiqiao Meng, Xiangnan He, Joemon Jose, Fuli Feng.
Accepted to WWW 2022. (Accept rate: 17.7%)

Improving Out-of-Distribution Robustness via Selective Augmentation. PDF
Huaxiu Yao*, **Yu Wang***, Sai Li, Weixin Liang, Linjun Zhang, James Zou, Chelsea Finn
Accepted to ICML 2022. (Accept rate: 21.9%)

Differentiable Invariant Causal Discovery. PDF
Yu Wang, An Zhang, Xiang Wang, Xiangnan He, Tat-Seng Chua
Submitted to NeurIPS 2022.

Interpretable Outlier Summarization
Yu Wang, Lei Cao, Sammuel Madden
Going to be submitted to PVLDB 2022.

Probabilistic and Variational Label Denoising.
Xin Xin*, **Yu Wang***, Zaiqiao Meng, Xiangnan He, Joemon Jose, Fuli Feng.
Going to be submitted to TOIS.

AutoOD: Automatic Outlier Detection.
Lei Cao, Yizhou Yan, **Yu Wang**, Samuel Madden, Elke A. Rundensteiner
Submitted to SIGMOD 2023.

RESEARCH EXPERIENCE

Topic: Invariant Causal Discovery *Sept. 2021 ∼ Oct. 2021*
Instructor: Prof. Tat-Seng Chua, Dr. Xiang Wang; NExT++, NUS
Goal: Utilize dataset from multiple environments to extract the invariant causal correlations.
- Proposed Invariant Causal Discovery(ICD) to enable robustness among multiple environments.
- Empirical results show that ICD achieves improvements up to 45% and 35% over the state-of-the-art baselines in linear and nonlinear setting, respectively.

Topic: Automatic Outlier Detection *Apr. 2021 ∼ Nov. 2021*
Instructor: Prof. Samuel Madden, Dr. Lei Cao; CSAIL, MIT
Goal: Detect the outliers with budgeted human evaluation resources.
- Proposed a new semi-supervised framework to fully utilize the limited human-labeled data.

[1] The mark "*" on the names means equal contribution.

- Developed a novel algorithm to generate interpretable and simple outlier summarizations.

Topic: Domain Generalization *Aug. 2021 ∼ Oct. 2021*
Instructor: Dr. Huaxiu Yao; SAIL, Stanford
Goal: Proposed the selective mixup strategy to handle the domain generalization problem.
- Proposed a novel mixup strategy for cancelling out spurious relations.
- Conducted enormous experiments on nine benchmark datasets ranging diverse domains, which demonstrate the superiority of our methods over seven previous methods.

Topic: Meta-Learning with an Adaptive Task Scheduler *Feb. 2021 ∼ May 2021*
Instructor: Dr. Huaxiu Yao; SAIL, Stanford
Goal: Construct a task sampling scheduler to deal with the meta-learning problems.
- Proposed a novel adaptive task scheduler for meta-learning.
- Conducted large-scale experiments on both an image classification benchmark (up to 13% improvement) and a real-world drug discovery dataset (up to 18% improvement).

Topic: Denoising in Recommendation *Oct. 2020 ∼ Feb. 2021*
Instructor: Prof. Xiangnan He, Dr. Xin Xin; Lab for Data Science, USTC
Goal: Debias(or denoise) in recommendation systems.
- Found the differences between the predictions of different models as the denoising signals.
- Proposed two methods to denoise the recommendation datasets in both interacted and uninteracted user-item pairs, achieving significant improvements.

PROJECT WORK

USTC-QA-System Github *Oct. 2020 ∼ Dec. 2020*
Instructor: Zhen-Hua Ling; Course: Natural Language Processing
- Collected about 200 questions specifically regarding USTC.
- Implemented a QA-system based on AskMSR(Brill et. al, EMNLP 2002) which achieved 51% accuracy on the collected test questions.

Image Segmentation Enhanced Style Transfer Github *Oct. 2020 ∼ Dec.2020*
Instructor: Yang Cao; Course: Computer Vision
- Proposed a novel framework to incorporate Image Segmentation into Style Transfer.
- Evaluated our framework based on CycleGAN and FastFCN and achieved fantastic results.

CityBrain Challenge Report *May 2021 ∼ June 2021*
Instructor: Defu Lian; Course:Introduction to Deep Learning
- Used rule-based methods to get a 24-th position in the challenge.
- Tried different Reinforcement Learning methods, and achieved the similar performance.

AWARDS AND HONORS

- Rose Fund New Lotus Scholarship (For outstanding research achievements). USTC, 2021.
- Baosteel Scholarship (One out of 402 students in School of the Gifted Young). USTC, 2021.
- Rose Fund Public Affairs Scholarship (For active leadership). USTC, 2021.
- Cyrus Tang Foundation Moral Education Scholarship. USTC, 2020.
- Huawei Scholarship (Top 3% of class). USTC, 2020.
- Excellent Student Scholarship – Gold (Top 3% of class). USTC, 2019.
- Provincial first prize in College Mathematics Competition(At most Top 7%). Anhui Province, 2019.

TECHNICAL STRENGTHS

English Test: TOEFL iBT 103 (104) (R: 25 (29), L: 28 (23), S: 25 (25), W: 25 (27)); GRE 320 + 3.5
Computer Skills: C, MATLAB, LaTeX, R, mathematica, JavaScript, Python, PyTorch, Tensorflow.