Daochen (Frank) Zha

Email: daochen.zha@rice.edu Phone: (979) 721-1408 Personal Webpage: https://dczha.com/
LinkedIn: https://www.linkedin.com/in/daochen-zha/ Address: 6100 Main Street, Houston, TX, 77005

EDUCATION

Rice University, Houston Aug. 2021 – May 2023 (Expected)
Ph.D. in Computer Science
GPA: 4.00 / 4.00 Advisor: Dr. Xia (Ben) Hu

Texas A&M University, College Station Sept. 2018 – Aug. 2021 (Transferred to Rice)
Ph.D. in Computer Science
GPA: 4.00 / 4.00 Advisor: Dr. Xia (Ben) Hu

Wuhan University, Wuhan, China Sept. 2014 - May 2018
Bachelor in Computer Science
GPA: 3.92 / 4.00 Advisor: Dr. Chenliang Li

HIGHLIGHTS

I am an experienced machine learning (ML) researcher and engineer, **leading the efforts** of several popular Automated Machine Learning (AutoML) and Reinforcement Learning (RL) open-source projects with **>5,000 stars** in total on GitHub (https://github.com/daochenzha), including DouZero (RL for Chinese poker DouDizhu), RLCard (a toolkit for RL in card games), TODS (AutoML for time-series outlier detection), and AutoVideo (AutoML for video recognition), etc. I have rich industrial ML experiences, including large-scale recommender systems, distributed ML training, and optimization with RL, etc. I was a research engineering intern at Meta Platforms, Inc., working on recommendation models.

INDUSTRIAL EXPERIENCE

Research Engineering Intern at **Meta Platforms, Inc.** May 2022– Aug. 2022
Mentor: Dr. Louis Feng and Dr. Yuandong Tian

- Studying system optimization for large-scale recommender systems (Ads models).
- Building cost models with deep neural networks.
- Studying learning-based device placement of operations with reinforcement learning.

Research Engineering Intern at **Facebook, Inc.** May 2021– Aug. 2021
Mentor: Dr. Louis Feng

- Designed automated searching algorithms with reinforcement learning for embedding table sharding (partitioning a large number of embedding tables into hundreds of GPU devices). Achieved 46% speedup over the sharding algorithms in production (**published in KDD'22 and NeurIPS'22, lead author**).

Research Intern at **Seattle AI Lab, Kuai Inc.** May 2020– Aug. 2020
Mentor: Dr. Wenye Ma and Dr. Ji Liu

- Studied reinforcement learning in DouDizhu, a challenging poker game in China. Improved the AI system by more than 4% in terms of win rate (**published in ICML'21, lead author**).
- Studied exploration in procedurally-generated environments. Achieved state-of-the-art performance on hard-exploration environments (**published in ICLR'21, lead author**).

SKILLS

Machine Learning: Automated Machine Learning, Reinforcement Learning, Machine Learning System, Anomaly and Outlier Detection, Graph Neural Networks, Recommender System, Text Mining, etc.
Programming Language: Python, Java, C/C++, MATLAB, Shell Script, HTML, SQL, LaTeX, etc.
Framework / Tools: PyTorch, TensorFlow, Keras, Scikit-Learn, NumPy, Pandas, etc.

OPEN-SOURCE PROJECTS

DouZero: Mastering DouDiZhu with Self-Play Deep Reinforcement Learning (**> 2,900 stars, lead author**)

- Developed a state-of-the-art AI system for the most popular Chinese poker DouDiZhu.
- Media coverage at Towards Data Science, Jiqizhixin (Chinese), Liangziwei (Chinese), Jiangmen (Chinese). Published in ICML 2022. Available at https://github.com/kwai/DouZero

RLCard: A Toolkit for Reinforcement Leaning in Card Games (**> 1,800 stars, lead author**)

- Leading and coordinating a team of 10 PhD/master students to develop and maintain the project.
- Media coverage at Synced, Towards Data Science, Jiqizhixin (Chinese), Leifengwang (Chinese) Zhiyuan (Chinese). Published in IJCAI 2022. Available at https://github.com/datamllab/rlcard

AutoVideo: An Automated Video Recognition System (**> 200 stars, lead author**)

- Leading and coordinating a team of 9 PhD/master students to develop and maintain the project.
- Published in IJCAI 2022. Available at https://github.com/datamllab/autovideo

TODS: An Automated Time-series Outlier Detection System (**> 600 stars, lead author**)

- Co-leading a team of 12 PhD/master students to develop and maintain the project.
- Published in AAAI 2021. Available at https://github.com/datamllab/tods

PUBLICATIONS

Conference (* means equal contribution)

NeurIPS'22 **Daochen Zha**, Louis Feng, Qiaoyu Tan, Zirui Liu, Kwei-Herng Lai, Bhargav Bhushanam, Yuandong Tian, Arun Kejariwal, Xia Hu. DreamShard: Generalizable Embedding Table Placement for Recommender Systems. Neural Information Processing Systems, 2022

CIKM'22 **Daochen Zha**, Kwei Herng Lai, Qiaoyu Tan, Sirui Ding, Na Zou, Xia Hu. Towards Automated Imbalanced Learning with Deep Hierarchical Reinforcement Learning. ACM International Conference on Information and Knowledge Management, 2022.

CIKM'22 **Daochen Zha*,** Guanchu Wang*, Zaid Pervaiz Bhat*, Zhimeng Jiang*, Yi-Wei Chen*, Alfredo Costilla Reyes*, Afshin Niktash, Gorkem Ulkar, Erman Okman, Xuanting Cai and Xia Hu. ACM International Conference on Information and Knowledge Management, demo track, 2022.

KDD'22 **Daochen Zha**, Louis Feng, Bhargav Bhushanam, Dhruv Choudhary, Jade Nie, Yuandong Tian, Jay Chae, Yinbin Ma, Arun Kejariwal, Xia Hu. AutoShard: Automated Embedding Table Sharding for Recommender Systems. ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2022.

KDD'22 Yuening Li, Zhengzhang Chen, **Daochen Zha**, Mengnan Du, Jingchao Ni, Denghui Zhang, Haifeng Chen, Xia Hu. Towards Learning Disentangled Representations for Time Series. ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2022.

IJCAI'22 **Daochen Zha***, Zaid Pervaiz Bhat*, Yi-Wei Chen*, Yicheng Wang*, Sirui Ding*, Jiaben Chen*, Kwei-Herng Lai*, Mohammad Qazim Bhat*, Anmoll Jain, Alfredo Costilla Reyes, Na Zou, Xia Hu. AutoVideo: An Automated Video Action Recognition System. International Joint Conference on Artificial Intelligence, demo track, 2022.

SDM'22 **Daochen Zha**, Kwei-Herng Lai, Kaixiong Zhou, Xia Hu. Towards Similarity-Aware Time-Series Classification. SIAM International Conference on Data Mining, 2022.

ICML'21 **Daochen Zha**, Jingru Xie, Wenye Ma, Sheng Zhang, Xiangru Lian, Xia Hu, Ji Liu. DouZero: Mastering DouDizhu with Self-Play Deep Reinforcement Learning. International Conference on Machine Learning, 2021.

ICLR'21 **Daochen Zha**, Wenye Ma, Lei Yuan, Xia Hu, Ji Liu. Rank the Episodes: A Simple Approach for Exploration in Procedurally-Generated Environments. International Conference on Learning Representations, 2021.

NeurIPS'21 Kaixiong Zhou, Xiao Huang, **Daochen Zha**, Rui Chen, Li Li, Soo-Hyun Choi, Xia Hu. Dirichlet Energy Constrained Learning for Deep Graph Neural Networks. Neural Information Processing Systems, 2021.

NeurIPS'21 Kwei-Herng Lai, **Daochen Zha**, Junjie Xu, Yue Zhao, Guanchu Wang, Xia Hu. Revisiting Time Series Outlier Detection: Definitions and Benchmarks. Neural Information Processing Systems, datasets and benchmarks track, 2021.

AAAI'21 **Daochen Zha***, Kwei-Herng Lai*, Guanchu Wang, Junjie Xu, Yue Zhao, Devesh Kumar, Yile Chen, Purav Zumkhawaka, Minyang Wan, Diego Martinez, Xia Hu. TODS: An Automated Time Series Outlier Detection System. AAAI Conference on Artificial Intelligence, demo track, 2021.

ICDE'21 Yuening Li, Zhengzhang Chen, **Daochen Zha**, Kaixiong Zhou, Haifeng Jin, Haifeng Chen, and Xia Hu. AutoOD: Neural Architecture Search for Outlier Detection. IEEE International Conference on Data Engineering, 2021.

ICDM'20 **Daochen Zha**, Kwei-Herng Lai, Mingyang Wan, and Xia Hu. Meta-AAD: Active Anomaly Detection with Deep Reinforcement Learning. IEEE International Conference on Data Mining, 2020.

IJCAI'20 **Daochen Zha*,** Kwei-Herng Lai*, Songyi Huang*, Yuanpu Cao, Keerthana Reddy, Juan Vargas, Alex Nguyen, Ruzhe Wei, Junyu Guo, Xia Hu. RLCard: A Platform for Reinforcement Learning in Card Games. International Joint Conference on Artificial Intelligence, demo track, 2020.

IJCAI'20 **Daochen Zha***, Kwei-Herng Lai*, Yuening Li, Xia Hu. Dual Policy Distillation. International Joint Conference on Artificial Intelligence, 2020.

IJCAI'20 Kaixiong Zhou, Qingquan Song, Xiao Huang, **Daochen Zha**, Na Zou, Xia Hu. Multi-Channel Graph Convolutional Networks. International Joint Conference on Artificial Intelligence, 2020.

KDD'20 Kwei Herng Lai, **Daochen Zha**, Kaixiong Zhou, and Xia Hu. PolicyGNN: Aggregation Optimization for Graph Neural Networks. ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2020.

WWW'20 Yuening Li, **Daochen Zha**, Na Zou, Xia Hu. PyODDS: An End-to-End Outlier Detection System. The Web Conference, demo track, 2020.

NeurIPS'20 Kaixiong Zhou, Xiao Huang, Yuening Li, **Daochen Zha**, Rui Chen, and Xia Hu. Towards Deeper Graph Neural Networks with Differentiable Group Normalization. Neural Information Processing Systems, 2020.

IJCAI'19 **Daochen Zha**, Kwei-Herng Lai, Kaixiong Zhou, Xia Hu. Experience Replay Optimization. International Joint Conference on Artificial Intelligence, 2019.

SCN'17 Yuan Shi, Huanguo Zhang, Juan Wang, Fei Yan, Bo Zhao, Feng Xiao, Jianwei Huang, **Daochen Zha**, Hongxin Hu. CHAOS: An SDN-Based Moving Target Defense System. Security and Communication Networks, 2017.

Journal

TKDD'22 Mingyang Wan, **Daochen Zha**, Ninghao Liu, Na Zou. In-Processing Modeling Techniques for Machine Learning Fairness: A Survey. ACM Transactions on Knowledge Discovery from Data, 2022.

TNNLS'21 Yuening Li, Zhengzhang Chen, **Daochen Zha**, Kaixiong Zhou, Haifeng Jin, Haifeng Chen, and Xia Hu. AutoOD: Automated Outlier Detection via Curiosity-guided Search and Self-imitation Learning. IEEE Transactions on Neural Networks and Learning Systems, 2021.

KAIS'19 **Daochen Zha**, Chenliang Li. Multi-label Dataless Text Classification with Topic Modeling. Knowledge and Information Systems, 2019.