

MINGHUAN LIU

Email: minghuanliu AT sjtu.edu.cn

Website: <http://minghuanliu.com>

EDUCATION

Shanghai Jiao Tong University (SJTU)

Sep. 2019 - Now

Ph.D. in Computer Science and Technology

- Apex Data & Knowledge Management Lab
- Leader of the **ApexRL** research group
- Advisor: Weinan Zhang

Southwest Jiaotong University (SWJTU)

Sep. 2015 - July. 2019

B.S. in Computer Science and Technology

- Overall GPA: 3.84/4.0 Ranking: 1/98
- Key Lab of Cloud Computing and Intelligent Technology
- Advisor: Tianrui Li

RESEARCH INTERESTS

- My general research interest lies in developing principled and efficient **reinforcement learning (RL)** algorithms to tackle kinds of decision making problems and build essential AI, including learning in **multi-agent systems**.

- I am sincerely devoted to **data-driven** RL methods, specifically, **imitation learning**, **offline RL**, that can make RL applicable for real-world challenges, like autonomous driving, sports analysis, healthcare, games, robotics, etc.

PUBLICATIONS / PREPRINTS

PerfectDou: Dominating DouDizhu with Perfect Information Distillation.

Guan Yang*, **Minghuan Liu***, Weijun Hong, Weinan Zhang, Fei Fang, Guangjun Zeng, Yue Lin.
(*Equal Contribution)

The 36th Conference on Neural Information Processing Systems. **NeurIPS 2022**.

- We propose PerfectDou, a state-of-the-art DouDizhu AI system that beats all previous algorithms with perfect information distillation and a perfect-training-imperfect-execution.

Reinforcement Learning with Automated Auxiliary Loss Search.

Tairan He, Yuge Zhang, Kan Ren, **Minghuan Liu**, Che Wang, Weinan Zhang, Dongsheng Li, Yuqing Yang.

The 36th Conference on Neural Information Processing Systems. **NeurIPS 2022**.

- We propose a principled and universal method for learning better representations with auxiliary loss functions, named Automated Auxiliary Loss Search (A2LS), which automatically searches for top-performing auxiliary loss functions for RL.

Plan Your Target and Learn Your Skills: Transferable State-Only Imitation Learning via Decoupled Policy Optimization.

Minghuan Liu, Zhengbang Zhu, Yuzheng Zhuang, Weinan Zhang, Jun Wang, Yong Yu, Jianye Hao.

The 39th International Conference on Machine Learning. **ICML 2022**.

- We propose Decoupled Policy Optimization (DePO), a novel framework that explicitly decouples the state-to-action mapping policy as a high-level satte planner and an inverse dynamics model. DePO allows for transferring to decision problems with different dynamics settings intuitively and generalizing the planner on out-of-demo state region.

Goal-Conditioned Reinforcement Learning: Problems and Solutions.

Minghuan Liu, Menghui Zhu, Weinan Zhang.

The 31st International Joint Conference on Artificial Intelligence, Survey Track. **IJCAI 2022**.

- We bring a brief survey for goal-conditioned reinforcement learning, containing the basic challenge, corresponding solutions, and future prospects.

Curriculum Offline Imitation Learning.

Minghuan Liu, Hanye Zhao, Zhengyu Yang, Jian Shen, Weinan Zhang, Li Zhao, Tie-Yan Liu.

The 35th Conference on Neural Information Processing Systems. **NeurIPS 2021**.

- We propose curriculum offline imitation learning (COIL), a simple and practical imitation learning based method for offline reinforcement learning. COIL utilizes an experience picking strategy for imitating from adaptive neighboring policies with a higher return, and improves the current policy along curriculum stages.

MapGo: Model-Assisted Policy Optimization for Goal-Oriented Tasks.

Menghui Zhu*, **Minghuan Liu***, Jian Shen, Zhicheng Zhang, Sheng Chen, Weinan Zhang, Deheng Ye, Yong Yu, Qiang Fu, Wei Yang. (*Equal Contribution)

The 30th International Joint Conference on Artificial Intelligence. **IJCAI 2021**.

- We propose MapGo, a model-based framework for goal-oriented RL which involves a novel relabeling strategy FGI and a model-based training module UMPO.

Energy-Based Imitation Learning.

Minghuan Liu, Tairan He, Minkai Xu, Weinan Zhang.

The 20th International Conference on Autonomous Agents and Multiagent Systems. **AAMAS 2021**.

- We propose EBIL, a two-step solution for imitation learning: first estimate the energy of expert's occupancy measure, and then take the energy to construct a surrogate reward function as a guidance for the agent to learn the desired policy.

Multi-Agent Interactions Modeling with Correlated Policies.

Minghuan Liu, Ming Zhou, Weinan Zhang, Yuzheng Zhuang, Jun Wang, Wulong Liu, Yong Yu.

The 8th International Conference on Learning Representations. **ICLR 2020**.

- We propose CoDAIL, which cast the multi-agent interactions modeling problem into a multi-agent imitation learning framework with explicit modeling of correlated policies by approximating opponents' policies.

Towards Applicable Reinforcement Learning: Improving the Generalization and Sample Efficiency with Policy Ensemble.

Zhengyu Yang, Kan Ren, Xufang Luo, **Minghuan Liu**, Weiqing Liu, Jiang Bian, Weinan Zhang, Dongsheng Li.

The 31st International Joint Conference on Artificial Intelligence **IJCAI 2022**.

- We propose Ensemble Proximal Policy Optimization (EPPO), a policy ensemble framework that improves the generalization and sample efficiency, especially in real-world tasks like financial trading.

Learning to Build High-fidelity and Robust Environment Models.

Weinan Zhang, Zhengyu Yang, Jian Shen, **Minghuan Liu**, Yimin Huang, Xing Zhang, Ruiming Tang, Zhenguo Li.

The 20th European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases. **ECML-PKDD 2021**.

- We propose robust learning to simulate (RL2S), a new problem of RL which focuses on learning a high-fidelity environment simulator for serving diverse downstream tasks; we further transform RL2S as a novel robust imitation learning problem and propose efficient algorithms to solve it.

AWARDS & HONORS

TOP 1, Ubiquant Retro Snake Challenge	<i>2022</i>
TOP 6, Finalist of Sports Analytics Challenge (sponsored by PSG)	<i>2019</i>
TOP 10, SCADA Data Missing Repair Competition	<i>2019</i>

TOP 3, AI Challenger 2018 in Weather Forecasting	<i>2018</i>
Outstanding Graduate of Sichuan	<i>2018</i>
National First Prize, China Undergraduate Mathematical Contest in Modeling	<i>2017</i>
Meritorious Winner, Mathematical Contest In Modeling	<i>2017</i>
China National Scholarship $\times 2$ (1%)	<i>2016&2017</i>
Tang Lixin Scholarship (1‰)	<i>2017</i>
IBM Scholarship (1‰)	<i>2017</i>
Special Grade Comprehensive Scholarship $\times 4$ (1%)	<i>2016 - 2018</i>

SKILLS

Machine Learning: Pytorch, Tensorflow, Scikit-Learn, LightGBM

Programming Languages: Python, JavaScript, C / C++, Java, MATLAB

Standard Tests: CET-6(574), CET4(616)

Hobbies and Interests: Soccer, Swimming