Junwei Liao

jwliao.ai@gmail.com / junweiliao@stu.xjtu.edu.cn / jwliao-ai.github.io

EDUCATION

Xi'an Jiaotong University   Xi'an, China
Bachelor of Science in Artificial Intelligence (Qian Xuesen Honors Program)   *Sep 2021 - Present*
- Cumulative **GPA**: 3.9/4.3, **Average Score**: 93.41/100 (w/ 3 bonus points), **CET6**: 602
- **Awards:** Outstanding Student (2021 - 2023), HUAWEI scholarship 2023 (0.1%)

University of California, Berkeley   Berkeley, CA, USA
Visiting Student   *Aug 2023 - Dec 2023*
- Cumulative **GPA:** 4.0/4.0
- **Relevant Coursework**: Deep Reinforcement Learning, Decision Making & Control (Graduate) | Designing, Visualizing and Understanding Deep Neural Networks | Theoretical Statistics

RESEARCH EXPERIENCE

Deep Reinforcement Learning Research Assistant Intern   Sep 2023 - Apr 2024
Dept. of Computer Science and Technology, Tsinghua University   *Supervised by Ju Ren*
- Reproduced most classic RL algorithms including DDPG, TD3, SAC, REINFORCE, REPARAMETRIZE, MBRL w/ on-policy data collection, and Random Shooting/CEM, Exploration and CQL/IQL/AWAC w/ online Finetuning.
- Reproduced classic RLAIF paper: Constitutional AI, where I collected AI feedback datasets and finetuned a pretrained OPT-1.3B using RL with AI feedback. Used vLLM and parallel computing to accelerate inference for dataset collection.
- Developed FeDPO, which combines Federated Learning and Direct Preference Optimization (a PbRL algorithm) with Ensemble-Directed Learning, where server can learn the optimal policy using pre-collected data from heterogeneous policies at clients w/o sharing the data.
- Investigated why in a small dataset (10K) setting, the performance of CQL/TD3 agent trained on expert dataset is much worse than that trained on medium-expert/random dataset, and visualized the trajectories of the original datasets and policy-collected trajectories using tSNE.

PROJECTS (DETAILS ON [GITHUB](https://github.com))

Online Diffuser: Planning with Diffusion Model   Mar 2024 - Present
- Designed a Diffusion agent to generate reasonable trajectories with high cumulative reward and be able to plan in the environment when given the initial state and target.
- Implemented Energy-Based Diffusion Model to explore and do adversarial learning to help our diffuser to generate diverse trajectories and ensure the validity of the synthetic dynamics/transitions and trajectories by compositional generation.
- Trying to design an effective value function to facilitate our diffuser to continuously learn from the feedback.

FeDPO: Federated Direct Preference Optimization   Jan 2024 - Mar 2024
- Reproduced classic RLHF algorithm DPO which dumps reward model learning and optimizes the policy directly by the concrete relationship between learning reward model and optimizing policy.
- Implemented DPO in a federated learning setting, used entropy regularization to prevent the weights from collapsing over a few clients, and utilized the performance of final local policy to achieve ensenble-directed learning.
- Leveraged FSDP (fully sharded data parallel) to significantly speed up training using multiple GPUs.

Neural ODE for Dynamics Learning in Continuous-Time MBRL   Oct 2023 - Dec 2023
- Implemented Vanilla/RNN/Augmented/Latent Neural ODE agents to learn dynamics in continuous-time MBRL using JAX.
- Considered continuous timestep governed by and ordinary differential equation, used NODE to model the instantaneous state change of a system and solved states by integrating the state changes.
- Conducted experiments in both on-policy and off-policy settings, and our best Neural ODE agent beat the baseline.
- Rated 100/100 for its novelty, scope, analysis and completeness in the CS285 final project session at UC Berkeley.

Evaluating LLMs with Psychometrics   Dec 2023
- Finetuned the LLM (using GPT-2 as base model) to imitate a variety of data (text styles), including Tweets, Reddit comments and news articles and observed how the psychometric results change.
- Implemented LoRA to achieve efficient finetuning, explored a series of bottleneck-sizes in our LoRA configuration, conducted an in-depth analysis of style-imitation performance and compute.
- Rated Excellent in the CS182 final project session at UC Berkeley.

Transformer-based Semantic Segmentation for LC Surgery   Oct 2022 - May 2023
- Played a key role in data augmentation and preprocessing in this project. Researched and implemented various data augmentation/generation techniques, including GAN, dynamic data augmentation to expand effective medical dataset.
- Investigated the use of Refine-Net, Vision Transformer and U-Net in semantic segmentation, and developed innovative TransUNet for the task. The multi-organ recognition accuracy reached 84.32%.
- Special Prize (top 2) in the Tengfei Cup at Xi'an Jiaotong University, and a related patent is being processed.