

Deep Reinforcement Learning

Course Project Proposal of Team 1

Tonghe Zhang^{1,*}, Yucheng Pan^{2,*}, Yulong Xia^{3,*}, Wei Zhang^{3,*}

¹ Department of Electronic Engineering

² Department of Physics

³ Department of Computer Science and Technology

{zhang-th21, panyuche21, xiayl23, zhangw23}@mails.tsinghua.edu.cn

2024-05-07

1 Topic

Project Objective Our course project is dedicated to exploring the offline RL track. The primary goal is to deepen our understanding of the established methodologies and cutting-edge solutions that enhance the performance of offline RL algorithms using limited data samples, as well as identifying the key factors that affect the generalizability and robustness of offline RL algorithms.

Research Focus A significant part of our investigation will explore how the composition, diversity, and volume of datasets influence the learning capabilities and performance of general offline RL algorithms[1][2][3]. Additionally, we are interested in examining whether data generation can effectively improve exploration and prevent over-fitting in offline RL settings.

Expected Contributions Our project aims to provide insights into the intricate relationship between the structure of the dataset and general offline RL algorithms. We also try to provide deeper understanding into the curious problem of whether synthesizing data using existing information deteriorates or enhances learning performance, which could serve as a corner-stone for the research in large-scale data augmentation in offline RL, sim2real RL and imitation learning.

2 Schedule

We tend to split our research project into three phases:

- **Week 10 to Week 11:** We will survey classical offline RL algorithms with trimming/augmentation techniques. We will also study relevant works concerning with exploration, data augmentation and over fitting, as well as “curiosity-driven RL” and “reward modification”.
- **Week 12-13:** We will implement different strategies to improve the offline RL dataset in parallel, meaning that each collaborator will in charge of one to two specific tasks.
- **Week 14-15:** We will investigate the relationship between exploration and data-generation, test whether reward-regularization during data-generation helps to yield better samples, and study whether introducing a dynamics model is necessary for offline RL tasks.

References

- [1] Guanghe Li, Yixiang Shan, Zhengbang Zhu, Ting Long, and Weinan Zhang. Diffstitch: Boosting offline reinforcement learning with diffusion-based trajectory stitching, 2024.
- [2] Cong Lu, Philip J. Ball, Yee Whye Teh, and Jack Parker-Holder. Synthetic experience replay, 2023.
- [3] Junjie Zhang, Jiafei Lyu, Xiaoteng Ma, Jiangpeng Yan, Jun Yang, Le Wan, and Xiu Li. Uncertainty-driven trajectory truncation for data augmentation in offline reinforcement learning, 2023.

¹ These authors contributed equally.