
awesome-offline-rl

This is a collection of research and review papers for **offline reinforcement learning (offline rl)**. Feel free to star and fork.

Maintainers: - Haruka Kiyohara (Cornell University) - Yuta Saito (Hanjuku-kaso Co., Ltd. / Cornell University)

We are looking for more contributors and maintainers! Please feel free to pull requests.

```
1 format:
2 - [title](paper link) [links]
3   - author1, author2, and author3. arXiv/conferences/journals/, year.
```

For any questions, feel free to contact: hk844@cornell.edu

Table of Contents

- Papers
 - Review/Survey/Position Papers
 - * Offline RL
 - * Off-Policy Evaluation and Learning
 - * Related Reviews
 - Offline RL: Theory/Methods
 - Offline RL: Benchmarks/Experiments
 - Offline RL: Applications
 - Off-Policy Evaluation and Learning: Theory/Methods
 - * Off-Policy Evaluation: Contextual Bandits
 - * Off-Policy Evaluation: Reinforcement Learning
 - * Off-Policy Learning
 - Off-Policy Evaluation and Learning: Benchmarks/Experiments
 - Off-Policy Evaluation and Learning: Applications
- Open Source Software/Implementations
- Blog/Podcast
 - Blog
 - Podcast
- Related Workshops
- Tutorials/Talks/Lectures

Papers

Review/Survey/Position Papers

Offline RL

- Open Problems and Fundamental Limitations of Reinforcement Learning from Human Feedback
 - Stephen Casper, Xander Davies, Claudia Shi, Thomas Krendl Gilbert, Jérémy Scheurer, Javier Rando, Rachel Freedman, Tomasz Korbak, David Lindner, Pedro Freire, Tony Wang, Samuel Marks, Charbel-Raphaël Segerie, Micah Carroll, Andi Peng, Phillip Christoffersen, Mehul Damani, Stewart Slocum, Usman Anwar, Anand Siththaranjan, Max Nadeau, Eric J. Michaud, Jacob Pfau, Dmitrii Krasheninnikov, Xin Chen, Lauro Langosco, Peter Hase, Erdem Biyik, Anca Dragan, David Krueger, Dorsa Sadigh, and Dylan Hadfield-Menell. arXiv, 2023.
- A Survey on Offline Model-Based Reinforcement Learning
 - Haoyang He. arXiv, 2023.
- Foundation Models for Decision Making: Problems, Methods, and Opportunities
 - Sherry Yang, Ofir Nachum, Yilun Du, Jason Wei, Pieter Abbeel, Dale Schuurmans. arXiv, 2023.
- A Survey on Offline Reinforcement Learning: Taxonomy, Review, and Open Problems
 - Rafael Figueiredo Prudencio, Marcos R. O. A. Maximo, and Esther Luna Colombini. arXiv, 2022.
- Offline Reinforcement Learning: Tutorial, Review, and Perspectives on Open Problems
 - Sergey Levine, Aviral Kumar, George Tucker, and Justin Fu. arXiv, 2020.

Off-Policy Evaluation and Learning

- A Review of Off-Policy Evaluation in Reinforcement Learning
 - Masatoshi Uehara, Chengchun Shi, and Nathan Kallus. arXiv, 2022.

Related Reviews

- On the Opportunities and Challenges of Offline Reinforcement Learning for Recommender Systems

-
- Xiaocong Chen, Siyu Wang, Julian McAuley, Dietmar Jannach, and Lina Yao. arXiv, 2023.
 - Understanding Reinforcement Learning Algorithms: The Progress from Basic Q-learning to Proximal Policy Optimization
 - Mohamed-Amine Chadi and Hajar Mousannif. arXiv, 2023.
 - Offline Evaluation for Reinforcement Learning-based Recommendation: A Critical Issue and Some Alternatives
 - Romain Deffayet, Thibaut Thonet, Jean-Michel Renders, and Maarten de Rijke. arXiv, 2023.
 - A Survey on Transformers in Reinforcement Learning
 - Wenzhe Li, Hao Luo, Zichuan Lin, Chongjie Zhang, Zongqing Lu, and Deheng Ye. arXiv, 2023.
 - Deep Reinforcement Learning: Opportunities and Challenges
 - Yuxi Li. arXiv, 2022.
 - A Survey on Model-based Reinforcement Learning
 - Fan-Ming Luo, Tian Xu, Hang Lai, Xiong-Hui Chen, Weinan Zhang, and Yang Yu. arXiv, 2022.
 - Survey on Fair Reinforcement Learning: Theory and Practice
 - Pratik Gajane, Akarti Saxena, Maryam Tavakol, George Fletcher, and Mykola Pechenizkiy. arXiv, 2022.
 - Accelerating Offline Reinforcement Learning Application in Real-Time Bidding and Recommendation: Potential Use of Simulation
 - Haruka Kiyohara, Kosuke Kawakami, and Yuta Saito. arXiv, 2021.
 - A Survey of Generalisation in Deep Reinforcement Learning
 - Robert Kirk, Amy Zhang, Edward Grefenstette, and Tim Rocktäschel. arXiv, 2021.

Offline RL: Theory/Methods

- Value-Aided Conditional Supervised Learning for Offline RL
 - Jeonghye Kim, Suyoung Lee, Woojun Kim, and Youngchul Sung. arXiv, 2024.
- Towards an Information Theoretic Framework of Context-Based Offline Meta-Reinforcement Learning

-
- Lanqing Li, Hai Zhang, Xinyu Zhang, Shatong Zhu, Junqiao Zhao, and Pheng-Ann Heng. arXiv, 2024.
 - DiffStitch: Boosting Offline Reinforcement Learning with Diffusion-based Trajectory Stitching
 - Guanghe Li, Yixiang Shan, Zhengbang Zhu, Ting Long, and Weinan Zhang. arXiv, 2024.
 - Deep autoregressive density nets vs neural ensembles for model-based offline reinforcement learning
 - Abdelhakim Benechehab, Albert Thomas, and Balázs Kégl. arXiv, 2024.
 - Context-Former: Stitching via Latent Conditioned Sequence Modeling
 - Ziqi Zhang, Jingzehua Xu, Zifeng Zhuang, Jinxin Liu, and Donglin wang. arXiv, 2024.
 - Adversarially Trained Actor Critic for offline CMDPs
 - Honghao Wei, Xiyue Peng, Xin Liu, and Arnob Ghosh. arXiv, 2024.
 - Optimistic Model Rollouts for Pessimistic Offline Policy Optimization
 - Yuanzhao Zhai, Yiyang Li, Zijian Gao, Xudong Gong, Kele Xu, Dawei Feng, Ding Bo, and Huaimin Wang. arXiv, 2024.
 - Solving Continual Offline Reinforcement Learning with Decision Transformer
 - Kaixin Huang, Li Shen, Chen Zhao, Chun Yuan, and Dacheng Tao. arXiv, 2024.
 - MoMA: Model-based Mirror Ascent for Offline Reinforcement Learning
 - Mao Hong, Zhiyue Zhang, Yue Wu, and Yanxun Xu. arXiv, 2024.
 - Reframing Offline Reinforcement Learning as a Regression Problem
 - Prajwal Koirala and Cody Fleming. arXiv, 2024.
 - Efficient Two-Phase Offline Deep Reinforcement Learning from Preference Feedback
 - Yinglun Xu and Gagandeep Singh. arXiv, 2024.
 - Policy-regularized Offline Multi-objective Reinforcement Learning
 - Qian Lin, Chao Yu, Zongkai Liu, and Zifan Wu. arXiv, 2024.
 - Differentiable Tree Search in Latent State Space
 - Dixant Mittal and Wee Sun Lee. arXiv, 2024.
 - Learning from Sparse Offline Datasets via Conservative Density Estimation
-

-
- Zhepeng Cen, Zuxin Liu, Zitong Wang, Yihang Yao, Henry Lam, and Ding Zhao. ICLR, 2024.
 - Safe Offline Reinforcement Learning with Feasibility-Guided Diffusion Model
 - Yinan Zheng, Jianxiong Li, Dongjie Yu, Yujie Yang, Shengbo Eben Li, Xianyuan Zhan, and Jingjing Liu. ICLR, 2024.
 - PDiT: Interleaving Perception and Decision-making Transformers for Deep Reinforcement Learning
 - Hangyu Mao, Rui Zhao, Ziyue Li, Zhiwei Xu, Hao Chen, Yiqun Chen, Bin Zhang, Zhen Xiao, Junge Zhang, and Jiangjin Yin. AAMAS, 2024.
 - Critic-Guided Decision Transformer for Offline Reinforcement Learning
 - Yuanfu Wang, Chao Yang, Ying Wen, Yu Liu, and Yu Qiao. AAAI, 2024.
 - CUDC: A Curiosity-Driven Unsupervised Data Collection Method with Adaptive Temporal Distances for Offline Reinforcement Learning
 - Chenyu Sun, Hangwei Qian, and Chunyan Miao. AAAI, 2024.
 - Neural Network Approximation for Pessimistic Offline Reinforcement Learning
 - Di Wu, Yuling Jiao, Li Shen, Haizhao Yang, and Xiliang Lu. AAAI, 2024.
 - A Perspective of Q-value Estimation on Offline-to-Online Reinforcement Learning
 - Yinmin Zhang, Jie Liu, Chuming Li, Yazhe Niu, Yaodong Yang, Yu Liu, and Wanli Ouyang. AAAI, 2024.
 - The Generalization Gap in Offline Reinforcement Learning
 - Ishita Mediratta, Qingfei You, Minqi Jiang, and Roberta Raileanu. arXiv, 2023.
 - Decoupling Meta-Reinforcement Learning with Gaussian Task Contexts and Skills
 - Hongcai He, Anjie Zhu, Shuang Liang, Feiyu Chen, and Jie Shao. arXiv, 2023.
 - MICRO: Model-Based Offline Reinforcement Learning with a Conservative Bellman Operator
 - Xiao-Yin Liu, Xiao-Hu Zhou, Guo-Tao Li, Hao Li, Mei-Jiang Gui, Tian-Yu Xiang, De-Xing Huang, and Zeng-Guang Hou. arXiv, 2023.
 - Model-Based Epistemic Variance of Values for Risk-Aware Policy Optimization
 - Carlos E. Luis, Alessandro G. Bottero, Julia Vinogradska, Felix Berkenkamp, and Jan Peters. arXiv, 2023.
 - Using Curiosity for an Even Representation of Tasks in Continual Offline Reinforcement Learning
-

-
- Pankayaraj Pathmanathan, Natalia Díaz-Rodríguez, and Javier Del Ser. arXiv, 2023.
 - Projected Off-Policy Q-Learning (POP-QL) for Stabilizing Offline Reinforcement Learning
 - Melrose Roderick, Gaurav Manek, Felix Berkenkamp, and J. Zico Kolter. arXiv, 2023.
 - Offline Data Enhanced On-Policy Policy Gradient with Provable Guarantees
 - Yifei Zhou, Ayush Sekhari, Yuda Song, and Wen Sun. arXiv, 2023.
 - Switch Trajectory Transformer with Distributional Value Approximation for Multi-Task Reinforcement Learning
 - Qinjie Lin, Han Liu, and Biswa Sengupta. arXiv, 2023.
 - Hierarchical Decision Transformer
 - André Correia and Luís A. Alexandre. arXiv, 2023.
 - Prompt-Tuning Decision Transformer with Preference Ranking
 - Shengchao Hu, Li Shen, Ya Zhang, and Dacheng Tao. arXiv, 2023.
 - Context Shift Reduction for Offline Meta-Reinforcement Learning
 - Yunkai Gao, Rui Zhang, Jiaming Guo, Fan Wu, Qi Yi, Shaohui Peng, Siming Lan, Ruizhi Chen, Zidong Du, Xing Hu, Qi Guo, Ling Li, and Yunji Chen. arXiv, 2023.
 - Uni-O4: Unifying Online and Offline Deep Reinforcement Learning with Multi-Step On-Policy Optimization
 - Kun Lei, Zhengmao He, Chenhao Lu, Kaizhe Hu, Yang Gao, and Huazhe Xu. arXiv, 2023.
 - Score Models for Offline Goal-Conditioned Reinforcement Learning
 - Harshit Sikchi, Rohan Chitnis, Ahmed Touati, Alborz Geramifard, Amy Zhang, and Scott Niekum. arXiv, 2023.
 - Offline RL with Observation Histories: Analyzing and Improving Sample Complexity
 - Joey Hong, Anca Dragan, and Sergey Levine. arXiv, 2023.
 - Expressive Modeling Is Insufficient for Offline RL: A Tractable Inference Perspective
 - Xuejie Liu, Anji Liu, Guy Van den Broeck, and Yitao Liang. arXiv, 2023.
 - Rethinking Decision Transformer via Hierarchical Reinforcement Learning
 - Yi Ma, Chenjun Xiao, Hebin Liang, and Jianye Hao. arXiv, 2023.
 - Unleashing the Power of Pre-trained Language Models for Offline Reinforcement Learning
-

-
- Ruizhe Shi, Yuyao Liu, Yanjie Ze, Simon S. Du, and Huazhe Xu. arXiv, 2023.
 - GOPlan: Goal-conditioned Offline Reinforcement Learning by Planning with Learned Models
 - Mianchu Wang, Rui Yang, Xi Chen, and Meng Fang. arXiv, 2023.
 - SERA: Sample Efficient Reward Augmentation in offline-to-online Reinforcement Learning
 - Ziqi Zhang, Xiao Xiong, Zifeng Zhuang, Jinxin Liu, and Donglin Wang. arXiv, 2023.
 - Bridging Distributionally Robust Learning and Offline RL: An Approach to Mitigate Distribution Shift and Partial Data Coverage
 - Kishan Panaganti, Zaiyan Xu, Dileep Kalathil, and Mohammad Ghavamzadeh. arXiv, 2023.
 - Guided Data Augmentation for Offline Reinforcement Learning and Imitation Learning
 - Nicholas E. Corrado, Yuxiao Qu, John U. Balis, Adam Labiosa, and Josiah P. Hanna. arXiv, 2023.
 - CROP: Conservative Reward for Model-based Offline Policy Optimization
 - Hao Li, Xiao-Hu Zhou, Xiao-Liang Xie, Shi-Qi Liu, Zhen-Qiu Feng, Xiao-Yin Liu, Mei-Jiang Gui, Tian-Yu Xiang, De-Xing Huang, Bo-Xian Yao, and Zeng-Guang Hou. arXiv, 2023.
 - Towards Robust Offline Reinforcement Learning under Diverse Data Corruption
 - Rui Yang, Han Zhong, Jiawei Xu, Amy Zhang, Chongjie Zhang, Lei Han, and Tong Zhang. arXiv, 2023.
 - Offline Retraining for Online RL: Decoupled Policy Learning to Mitigate Exploration Bias
 - Max Sobol Mark, Archit Sharma, Fahim Tajwar, Rafael Rafailov, Sergey Levine, and Chelsea Finn. arXiv, 2023.
 - Boosting Continuous Control with Consistency Policy
 - Yuhui Chen, Haoran Li, and Dongbin Zhao. arXiv, 2023.
 - Planning to Go Out-of-Distribution in Offline-to-Online Reinforcement Learning
 - Trevor McInroe, Stefano V. Albrecht, and Amos Storkey. arXiv, 2023.
 - Reward-Consistent Dynamics Models are Strongly Generalizable for Offline Reinforcement Learning
 - Fan-Ming Luo, Tian Xu, Xingchen Cao, and Yang Yu. arXiv, 2023.
 - DiffCPS: Diffusion Model based Constrained Policy Search for Offline Reinforcement Learning
-

-
- Longxiang He, Linrui Zhang, Junbo Tan, and Xueqian Wang. arXiv, 2023.
 - Self-Confirming Transformer for Locally Consistent Online Adaptation in Multi-Agent Reinforcement Learning
 - Tao Li, Juan Guevara, Xinghong Xie, and Quanyan Zhu. arXiv, 2023.
 - Learning to Reach Goals via Diffusion
 - Vineet Jain and Siamak Ravanbakhsh. arXiv, 2023.
 - Decision ConvFormer: Local Filtering in MetaFormer is Sufficient for Decision Making
 - Jeonghye Kim, Suyoung Lee, Woojun Kim, and Youngchul Sung. arXiv, 2023.
 - Consistency Models as a Rich and Efficient Policy Class for Reinforcement Learning
 - Zihan Ding and Chi Jin. arXiv, 2023.
 - Pessimistic Nonlinear Least-Squares Value Iteration for Offline Reinforcement Learning
 - Qiwei Di, Heyang Zhao, Jiafan He, and Quanquan Gu. arXiv, 2023.
 - Reasoning with Latent Diffusion in Offline Reinforcement Learning
 - Siddarth Venkatraman, Shivesh Khaitan, Ravi Tej Akella, John Dolan, Jeff Schneider, and Glen Berseth. arXiv, 2023.
 - Hundreds Guide Millions: Adaptive Offline Reinforcement Learning with Expert Guidance
 - Qisen Yang, Shenzhi Wang, Qihang Zhang, Gao Huang, and Shiji Song. arXiv, 2023.
 - Towards Robust Offline-to-Online Reinforcement Learning via Uncertainty and Smoothness
 - Xiaoyu Wen, Xudong Yu, Rui Yang, Chenjia Bai, and Zhen Wang. arXiv, 2023.
 - Robust Offline Reinforcement Learning – Certify the Confidence Interval
 - Jiarui Yao and Simon Shaolei Du. arXiv, 2023.
 - Stackelberg Batch Policy Learning
 - Wenzhuo Zhou and Annie Qu. arXiv, 2023.
 - H2O+: An Improved Framework for Hybrid Offline-and-Online RL with Dynamics Gaps
 - Haoyi Niu, Tianying Ji, Bingqi Liu, Haocheng Zhao, Xiangyu Zhu, Jianying Zheng, Pengfei Huang, Guyue Zhou, Jianming Hu, and Xianyu Zhan. arXiv, 2023.
 - Q-Transformer: Scalable Offline Reinforcement Learning via Autoregressive Q-Functions
-

-
- Yevgen Chebotar, Quan Vuong, Alex Irpan, Karol Hausman, Fei Xia, Yao Lu, Aviral Kumar, Tianhe Yu, Alexander Herzog, Karl Pertsch, Keerthana Gopalakrishnan, Julian Ibarz, Ofir Nachum, Sumedh Sontakke, Grecia Salazar, Huong T Tran, Jodilyn Peralta, Clayton Tan, Deeksha Manjunath, Jaspiar Singht, Brianna Zitkovich, Tomas Jackson, Kanishka Rao, Chelsea Finn, and Sergey Levine. arXiv, 2023.
 - DOMAIN: Mildly Conservative Model-BASED OFFLINE Reinforcement Learning
 - Xiao-Yin Liu, Xiao-Hu Zhou, Xiao-Liang Xie, Shi-Qi Liu, Zhen-Qiu Feng, Hao Li, Mei-Jiang Gui, Tian-Yu Xiang, De-Xing Huang, and Zeng-Guang Hou. arXiv, 2023.
 - Guided Online Distillation: Promoting Safe Reinforcement Learning by Offline Demonstration
 - Jinning Li, Xinyi Liu, Banghua Zhu, Jiantao Jiao, Masayoshi Tomizuka, Chen Tang, and Wei Zhan. arXiv, 2023.
 - Equivariant Data Augmentation for Generalization in Offline Reinforcement Learning
 - Cristina Pinneri, Sarah Bechtle, Markus Wulfmeier, Arunkumar Byravan, Jingwei Zhang, William F. Whitney, and Martin Riedmiller. arXiv, 2023.
 - Reasoning with Latent Diffusion in Offline Reinforcement Learning
 - Siddarth Venkatraman, Shivesh Khaitan, Ravi Tej Akella, John Dolan, Jeff Schneider, and Glen Berseth. arXiv, 2023.
 - Hundreds Guide Millions: Adaptive Offline Reinforcement Learning with Expert Guidance
 - Qisen Yang, Shenzhi Wang, Qihang Zhang, Gao Huang, and Shiji Song. arXiv, 2023.
 - Multi-Objective Decision Transformers for Offline Reinforcement Learning
 - Abdelghani Ghanem, Philippe Ciblat, and Mounir Ghogho. arXiv, 2023.
 - AlphaStar Unplugged: Large-Scale Offline Reinforcement Learning
 - Michaël Mathieu, Sherjil Ozair, Srivatsan Srinivasan, Caglar Gulcehre, Shangdong Zhang, Ray Jiang, Tom Le Paine, Richard Powell, Konrad Żoźna, Julian Schrittwieser, David Choi, Petko Georgiev, Daniel Toyama, Aja Huang, Roman Ring, Igor Babuschkin, Timo Ewalds, Mahyar Bordbar, Sarah Henderson, Sergio Gómez Colmenarejo, Aäron van den Oord, Wojciech Marian Czarnecki, Nando de Freitas, and Oriol Vinyals. arXiv, 2023.
 - Exploiting Generalization in Offline Reinforcement Learning via Unseen State Augmentations
 - Nirbhay Modhe, Qiaozi Gao, Ashwin Kalyan, Dhruv Batra, Govind Thattai, and Gaurav Sukhatme. arXiv, 2023.
-

-
- PASTA: Pretrained Action-State Transformer Agents
 - Raphael Boige, Yannis Flet-Berliac, Arthur Flajolet, Guillaume Richard, and Thomas Pierrot. arXiv, 2023.
 - Towards A Unified Agent with Foundation Models
 - Norman Di Palo, Arunkumar Byravan, Leonard Hasenclever, Markus Wulfmeier, Nicolas Heess, and Martin Riedmiller. arXiv, 2023.
 - Goal-Conditioned Predictive Coding as an Implicit Planner for Offline Reinforcement Learning
 - Zilai Zeng, Ce Zhang, Shijie Wang, and Chen Sun. arXiv, 2023.
 - Offline Reinforcement Learning with Imbalanced Datasets
 - Li Jiang, Sijie Chen, Jieli Qiu, Haoran Xu, Wai Kin Chan, and Zhao Ding. arXiv, 2023.
 - LLQL: Logistic Likelihood Q-Learning for Reinforcement Learning
 - Outongyi Lv, Bingxin Zhou, and Yu Guang Wang. arXiv, 2023.
 - Elastic Decision Transformer
 - Yueh-Hua Wu, Xiaolong Wang, and Masashi Hamaya. arXiv, 2023.
 - Prioritized Trajectory Replay: A Replay Memory for Data-driven Reinforcement Learning
 - Jinyi Liu, Yi Ma, Jianye Hao, Yujing Hu, Yan Zheng, Tangjie Lv, and Changjie Fan. arXiv, 2023.
 - Is RLHF More Difficult than Standard RL?
 - Yuanhao Wang, Qinghua Liu, and Chi Jin. arXiv, 2023.
 - Supervised Pretraining Can Learn In-Context Reinforcement Learning
 - Jonathan N. Lee, Annie Xie, Aldo Pacchiano, Yash Chandak, Chelsea Finn, Ofir Nachum, and Emma Brunskill. arXiv, 2023.
 - Fighting Uncertainty with Gradients: Offline Reinforcement Learning via Diffusion Score Matching
 - H.J. Terry Suh, Glen Chou, Hongkai Dai, Lujie Yang, Abhishek Gupta, and Russ Tedrake. arXiv, 2023.
 - Safe Reinforcement Learning with Dead-Ends Avoidance and Recovery
 - Xiao Zhang, Hai Zhang, Hongtu Zhou, Chang Huang, Di Zhang, Chen Ye, and Junqiao Zhao. arXiv, 2023.

-
- CLUE: Calibrated Latent Guidance for Offline Reinforcement Learning
 - Jinxin Liu, Lipeng Zu, Li He, and Donglin Wang. arXiv, 2023.
 - Harnessing Mixed Offline Reinforcement Learning Datasets via Trajectory Weighting
 - Zhang-Wei Hong, Pulkit Agrawal, Rémi Tachet des Combes, and Romain Laroche.
 - Beyond OOD State Actions: Supported Cross-Domain Offline Reinforcement Learning
 - Jinxin Liu, Ziqi Zhang, Zhenyu Wei, Zifeng Zhuang, Yachen Kang, Sibao Gai, and Donglin Wang. arXiv, 2023.
 - A Primal-Dual-Critic Algorithm for Offline Constrained Reinforcement Learning
 - Kihyuk Hong, Yuhang Li, and Ambuj Tewari. arXiv, 2023.
 - HIPODE: Enhancing Offline Reinforcement Learning with High-Quality Synthetic Data from a Policy-Decoupled Approach
 - Shixi Lian, Yi Ma, Jinyi Liu, Yan Zheng, and Zhaopeng Meng. arXiv, 2023.
 - Ensemble-based Offline-to-Online Reinforcement Learning: From Pessimistic Learning to Optimistic Exploration
 - Kai Zhao, Yi Ma, Jinyi Liu, Yan Zheng, and Zhaopeng Meng. arXiv, 2023.
 - In-Sample Policy Iteration for Offline Reinforcement Learning
 - Xiaohan Hu, Yi Ma, Chenjun Xiao, Yan Zheng, and Zhaopeng Meng. arXiv, 2023.
 - Instructed Diffuser with Temporal Condition Guidance for Offline Reinforcement Learning
 - Jifeng Hu, Yanchao Sun, Sili Huang, SiYuan Guo, Hechang Chen, Li Shen, Lichao Sun, Yi Chang, and Dacheng Tao. arXiv, 2023.
 - Offline Prioritized Experience Replay
 - Yang Yue, Bingyi Kang, Xiao Ma, Gao Huang, Shiji Song, and Shuicheng Yan. arXiv, 2023.
 - Delphic Offline Reinforcement Learning under Nonidentifiable Hidden Confounding
 - Alizée Pace, Hugo Yèche, Bernhard Schölkopf, Gunnar Rätsch, and Guy Tennenholtz. arXiv, 2023.
 - Offline Meta Reinforcement Learning with In-Distribution Online Adaptation
 - Jianhao Wang, Jin Zhang, Haozhe Jiang, Junyu Zhang, Liwei Wang, and Chongjie Zhang. arXiv, 2023.

-
- Diffusion Model is an Effective Planner and Data Synthesizer for Multi-Task Reinforcement Learning
 - Haoran He, Chenjia Bai, Kang Xu, Zhuoran Yang, Weinan Zhang, Dong Wang, Bin Zhao, and Xuelong Li. arXiv, 2023.
 - Reinforcement Learning with Human Feedback: Learning Dynamic Choices via Pessimism
 - Zihao Li, Zhuoran Yang, and Mengdi Wang. arXiv, 2023.
 - MADiff: Offline Multi-agent Learning with Diffusion Models
 - Zhengbang Zhu, Minghuan Liu, Liyuan Mao, Bingyi Kang, Minkai Xu, Yong Yu, Stefano Ermon, and Weinan Zhang. arXiv, 2023.
 - Provable Offline Reinforcement Learning with Human Feedback
 - Wenhao Zhan, Masatoshi Uehara, Nathan Kallus, Jason D. Lee, and Wen Sun. arXiv, 2023.
 - Think Before You Act: Decision Transformers with Internal Working Memory
 - Jikun Kang, Romain Laroche, Xindi Yuan, Adam Trischler, Xue Liu, and Jie Fu. arXiv, 2023.
 - Distributionally Robust Optimization Efficiently Solves Offline Reinforcement Learning
 - Yue Wang, Yuting Hu, Jinjun Xiong, and Shaofeng Zou. arXiv, 2023.
 - Offline Primal-Dual Reinforcement Learning for Linear MDPs
 - Germano Gabbianelli, Gergely Neu, Nneka Okolo, and Matteo Papini. arXiv, 2023.
 - Federated Offline Policy Learning with Heterogeneous Observational Data
 - Aldo Gael Carranza and Susan Athey. arXiv, 2023.
 - Offline Reinforcement Learning with Additional Covering Distributions
 - Chenjie Mao. arXiv, 2023.
 - Reward-agnostic Fine-tuning: Provable Statistical Benefits of Hybrid Reinforcement Learning
 - Gen Li, Wenhao Zhan, Jason D. Lee, Yuejie Chi, and Yuxin Chen. arXiv, 2023.
 - Stackelberg Decision Transformer for Asynchronous Action Coordination in Multi-Agent Systems
 - Bin Zhang, Hangyu Mao, Lijuan Li, Zhiwei Xu, Dapeng Li, Rui Zhao, and Guoliang Fan. arXiv, 2023.
 - Federated Ensemble-Directed Offline Reinforcement Learning

-
- Desik Rengarajan, Nitin Ragothaman, Dileep Kalathil, and Srinivas Shakkottai. arXiv, 2023.
 - IDQL: Implicit Q-Learning as an Actor-Critic Method with Diffusion Policies
 - Philippe Hansen-Estruch, Ilya Kostrikov, Michael Janner, Jakub Grudzien Kuba, and Sergey Levine. arXiv, 2023.
 - Using Offline Data to Speed-up Reinforcement Learning in Procedurally Generated Environments
 - Alain Andres, Lukas Schäfer, Esther Villar-Rodriguez, Stefano V.Albrecht, Javier Del Ser. arXiv, 2023.
 - Reinforcement Learning from Passive Data via Latent Intentions [website]
 - Dibya Ghosh, Chethan Bhateja, and Sergey Levine. arXiv, 2023.
 - Uncertainty-driven Trajectory Truncation for Model-based Offline Reinforcement Learning
 - Junjie Zhang, Jiafei Lyu, Xiaoteng Ma, Jiangpeng Yan, Jun Yang, Le Wan, and Xiu Li. arXiv, 2023.
 - RAFT: Reward rAnked FineTuning for Generative Foundation Model Alignment
 - Hanze Dong, Wei Xiong, Deepanshu Goyal, Rui Pan, Shizhe Diao, Jipeng Zhang, Kashun Shum, and Tong Zhang. arXiv, 2023.
 - Batch Quantum Reinforcement Learning
 - Maniraman Periyasamy, Marc Hölle, Marco Wiedmann, Daniel D. Scherer, Axel Plinge, and Christopher Mutschler. arXiv, 2023.
 - Accelerating exploration and representation learning with offline pre-training
 - Bogdan Mazouze, Jake Bruce, Doina Precup, Rob Fergus, and Ankit Anand. arXiv, 2023.
 - On Context Distribution Shift in Task Representation Learning for Offline Meta RL
 - Chenyang Zhao, Zihao Zhou, and Bin Liu. arXiv, 2023.
 - Optimal Goal-Reaching Reinforcement Learning via Quasimetric Learning
 - Tongzhou Wang, Antonio Torralba, Phillip Isola, and Amy Zhang. arXiv, 2023.
 - Learning Excavation of Rigid Objects with Offline Reinforcement Learning
 - Shiyu Jin, Zhixian Ye, and Liangjun Zhang. arXiv, 2023.
 - Goal-conditioned Offline Reinforcement Learning through State Space Partitioning
-

-
- Mianchu Wang, Yue Jin, and Giovanni Montana. arXiv, 2023.
 - Merging Decision Transformers: Weight Averaging for Forming Multi-Task Policies
 - Daniel Lawson and Ahmed H. Qureshi. arXiv, 2023.
 - Deploying Offline Reinforcement Learning with Human Feedback
 - Ziniu Li, Ke Xu, Liu Liu, Lanqing Li, Deheng Ye, and Peilin Zhao. arXiv, 2023.
 - Synthetic Experience Replay
 - Cong Lu, Philip J. Ball, and Jack Parker-Holder. arXiv, 2023.
 - ENTROPY: Environment Transformer and Offline Policy Optimization
 - Pengqin Wang, Meixin Zhu, and Shaojie Shen. arXiv, 2023.
 - Graph Decision Transformer
 - Shengchao Hu, Li Shen, Ya Zhang, and Dacheng Tao. arXiv, 2023.
 - Selective Uncertainty Propagation in Offline RL
 - Sanath Kumar Krishnamurthy, Tanmay Gangwani, Sumeet Katariya, Branislav Kveton, and Anshuka Rangi. arXiv, 2023.
 - Off-the-Grid MARL: a Framework for Dataset Generation with Baselines for Cooperative Offline Multi-Agent Reinforcement Learning
 - Claude Formanek, Asad Jeewa, Jonathan Shock, and Arnu Pretorius. arXiv, 2023.
 - Skill Decision Transformer
 - Shyam Sudhakaran and Sebastian Risi. arXiv, 2023.
 - Guiding Online Reinforcement Learning with Action-Free Offline Pretraining
 - Deyao Zhu, Yuhui Wang, Jürgen Schmidhuber, and Mohamed Elhoseiny. arXiv, 2023.
 - SaFormer: A Conditional Sequence Modeling Approach to Offline Safe Reinforcement Learning
 - Qin Zhang, Linrui Zhang, Haoran Xu, Li Shen, Bowen Wang, Yongzhe Chang, Xueqian Wang, Bo Yuan, and Dacheng Tao. arXiv, 2023.
 - APAC: Authorized Probability-controlled Actor-Critic For Offline Reinforcement Learning
 - Jing Zhang, Chi Zhang, Wenjia Wang, and Bing-Yi Jing. arXiv, 2023.
 - Designing an offline reinforcement learning objective from scratch
-

-
- Gaon An, Junhyeok Lee, Xingdong Zuo, Norio Kosaka, Kyung-Min Kim, and Hyun Oh Song. arXiv, 2023.
 - Behaviour Discriminator: A Simple Data Filtering Method to Improve Offline Policy Learning
 - Qiang Wang, Robert McCarthy, David Cordova Bulens, Kevin McGuinness, Noel E. O’Connor, Francisco Roldan Sanchez, and Stephen J. Redmond. arXiv, 2023.
 - Learning to View: Decision Transformers for Active Object Detection
 - Wenhao Ding, Nathalie Majcherczyk, Mohit Deshpande, Xuwei Qi, Ding Zhao, Rajasimman Madhivanan, and Arnie Sen. arXiv, 2023.
 - Risk Sensitive Dead-end Identification in Safety-Critical Offline Reinforcement Learning
 - Taylor W. Killian, Sonali Parbhoo, and Marzyeh Ghassemi. arXiv, 2023.
 - Value Enhancement of Reinforcement Learning via Efficient and Robust Trust Region Optimization
 - Chengchun Shi, Zhengling Qi, Jianing Wang, and Fan Zhou. arXiv, 2023.
 - Contextual Conservative Q-Learning for Offline Reinforcement Learning
 - Ke Jiang, Jiayu Yao, and Xiaoyang Tan. arXiv, 2023.
 - Offline Policy Optimization in RL with Variance Regularization
 - Riashat Islam, Samarth Sinha, Homanga Bharadhwaj, Samin Yeasar Arnob, Zhuoran Yang, Animesh Garg, Zhaoran Wang, Lihong Li, and Doina Precup. arXiv, 2023.
 - Transformer in Transformer as Backbone for Deep Reinforcement Learning
 - Hangyu Mao, Rui Zhao, Hao Chen, Jianye Hao, Yiqun Chen, Dong Li, Junge Zhang, and Zhen Xiao. arXiv, 2023.
 - SPQR: Controlling Q-ensemble Independence with Spiked Random Model for Reinforcement Learning
 - Dohyeok Lee, Seungyub Han, Taehyun Cho, and Jungwoo Lee. NeurIPS, 2023.
 - Revisiting the Minimalist Approach to Offline Reinforcement Learning
 - Denis Tarasov, Vladislav Kurenkov, Alexander Nikulin, and Sergey Kolesnikov. NeurIPS, 2023.
 - Constrained Policy Optimization with Explicit Behavior Density for Offline Reinforcement Learning
-

-
- Jing Zhang, Chi Zhang, Wenjia Wang, and Bingyi Jing. NeurIPS, 2023.
 - Supported Value Regularization for Offline Reinforcement Learning
 - Yixiu Mao, Hongchang Zhang, Chen Chen, Yi Xu, and Xiangyang Ji. NeurIPS, 2023.
 - Conservative State Value Estimation for Offline Reinforcement Learning
 - Liting Chen, Jie Yan, Zhengdao Shao, Lu Wang, Qingwei Lin, Saravan Rajmohan, Thomas Moscibroda, and Dongmei Zhang. NeurIPS, 2023.
 - Understanding and Addressing the Pitfalls of Bisimulation-based Representations in Offline Reinforcement Learning
 - Hongyu Zang, Xin Li, Leiji Zhang, Yang Liu, Baigui Sun, Riashat Islam, Remi Tachet des Combes, and Romain Laroche. NeurIPS, 2023.
 - Adversarial Model for Offline Reinforcement Learning
 - Mohak Bhardwaj, Tengyang Xie, Byron Boots, Nan Jiang, and Ching-An Cheng. NeurIPS, 2023.
 - Percentile Criterion Optimization in Offline Reinforcement Learning
 - Cyrus Cousins, Elita Lobo, Marek Petrik, and Yair Zick. NeurIPS, 2023.
 - Importance Weighted Actor-Critic for Optimal Conservative Offline Reinforcement Learning
 - Hanlin Zhu, Paria Rashidinejad, and Jiantao Jiao. NeurIPS, 2023.
 - HIQL: Offline Goal-Conditioned RL with Latent States as Actions
 - Seohong Park, Dibya Ghosh, Benjamin Eysenbach, and Sergey Levine. NeurIPS, 2023.
 - Recovering from Out-of-sample States via Inverse Dynamics in Offline Reinforcement Learning
 - Ke Jiang, Jia-Yu Yao, and Xiaoyang Tan. NeurIPS, 2023.
 - Offline RL with Discrete Proxy Representations for Generalizability in POMDPs
 - Pengjie Gu, Xinyu Cai, Dong Xing, Xinrun Wang, Mengchen Zhao, and Bo An. NeurIPS, 2023.
 - Offline Multi-Agent Reinforcement Learning with Implicit Global-to-Local Value Regularization
 - Xiangsen Wang, Haoran Xu, Yinan Zheng, and Xianyu Zhan. NeurIPS, 2023.
 - Bi-Level Offline Policy Optimization with Limited Exploration
 - Wenzhuo Zhou. NeurIPS, 2023.
 - Provably (More) Sample-Efficient Offline RL with Options

-
- Xiaoyan Hu and Ho-fung Leung. NeurIPS, 2023.
 - Double Pessimism is Provably Efficient for Distributionally Robust Offline Reinforcement Learning: Generic Algorithm and Robust Partial Coverage
 - Jose Blanchet, Miao Lu, Tong Zhang, and Han Zhong. NeurIPS, 2023.
 - AlberDICE: Addressing Out-Of-Distribution Joint Actions in Offline Multi-Agent RL via Alternating Stationary Distribution Correction Estimation
 - Daiki E. Matsunaga, Jongmin Lee, Jaeseok Yoon, Stefanos Leonardos, Pieter Abbeel, and Kee-Eung Kim. NeurIPS, 2023.
 - Budgeting Counterfactual for Offline RL
 - Yao Liu, Pratik Chaudhari, and Rasool Fakoor. NeurIPS, 2023.
 - Efficient Diffusion Policies for Offline Reinforcement Learning
 - Bingyi Kang, Xiao Ma, Chao Du, Tianyu Pang, and Shuicheng Yan. NeurIPS, 2023.
 - Cal-QL: Calibrated Offline RL Pre-Training for Efficient Online Fine-Tuning
 - Mitsuhiro Nakamoto, Yuexiang Zhai, Anikait Singh, Max Sobol Mark, Yi Ma, Chelsea Finn, Aviral Kumar, and Sergey Levine. NeurIPS, 2023.
 - Policy Finetuning in Reinforcement Learning via Design of Experiments using Offline Data
 - Ruiqi Zhang and Andrea Zanette. NeurIPS, 2023.
 - Offline Minimax Soft-Q-learning Under Realizability and Partial Coverage
 - Masatoshi Uehara, Nathan Kallus, Jason D. Lee, and Wen Sun. NeurIPS, 2023.
 - Provably Efficient Offline Reinforcement Learning in Regular Decision Processes
 - Roberto Cipollone, Anders Jonsson, Alessandro Ronca, and Mohammad Sadegh Talebi. NeurIPS, 2023.
 - Provably Efficient Offline Goal-Conditioned Reinforcement Learning with General Function Approximation and Single-Policy Concentrability
 - Hanlin Zhu and Amy Zhang. NeurIPS, 2023.
 - On Sample-Efficient Offline Reinforcement Learning: Data Diversity, Posterior Sampling and Beyond
 - Thanh Nguyen-Tang and Raman Arora. NeurIPS, 2023.
 - Conservative Offline Policy Adaptation in Multi-Agent Games
-

-
- Chengjie Wu, Pingzhong Tang, Jun Yang, Yujing Hu, Tangjie Lv, Changjie Fan, and Chongjie Zhang. NeurIPS, 2023.
 - Look Beneath the Surface: Exploiting Fundamental Symmetry for Sample-Efficient Offline RL
 - Peng Cheng, Xianyuan Zhan, Zhihao Wu, Wenjia Zhang, Shoucheng Song, Han Wang, Youfang Lin, and Li Jiang. NeurIPS, 2023.
 - Survival Instinct in Offline Reinforcement Learning
 - Anqi Li, Dipendra Misra, Andrey Kolobov, and Ching-An Cheng. NeurIPS, 2023.
 - Learning from Visual Observation via Offline Pretrained State-to-Go Transformer
 - Bohan Zhou, Ke Li, Jiechuan Jiang, and Zongqing Lu. NeurIPS, 2023.
 - Design from Policies: Conservative Test-Time Adaptation for Offline Policy Optimization
 - Jinxin Liu, Hongyin Zhang, Zifeng Zhuang, Yachen Kang, Donglin Wang, and Bin Wang. NeurIPS, 2023.
 - Learning to Influence Human Behavior with Offline Reinforcement Learning
 - Joey Hong, Anca Dragan, and Sergey Levine. NeurIPS, 2023.
 - Residual Q-Learning: Offline and Online Policy Customization without Value
 - Chenran Li, Chen Tang, Haruki Nishimura, Jean Mercat, Masayoshi Tomizuka, Wei Zhan. NeurIPS, 2023.
 - Train Once, Get a Family: State-Adaptive Balances for Offline-to-Online Reinforcement Learning
 - Shenzhi Wang, Qisen Yang, Jiawei Gao, Matthieu Gaetan Lin, Hao Chen, Liwei Wu, Ning Jia, Shiji Song, and Gao Huang. NeurIPS, 2023.
 - Beyond Uniform Sampling: Offline Reinforcement Learning with Imbalanced Datasets
 - Zhang-Wei Hong, Aviral Kumar, Sathwik Karnik, Abhishek Bhandwaldar, Akash Srivastava, Joni Pajarinen, Romain Laroche, Abhishek Gupta, and Pulkit Agrawal. NeurIPS, 2023.
 - Understanding, Predicting and Better Resolving Q-Value Divergence in Offline-RL
 - Yang Yue, Rui Lu, Bingyi Kang, Shiji Song, and Gao Huang. NeurIPS, 2023.
 - Corruption-Robust Offline Reinforcement Learning with General Function Approximation
 - Chenlu Ye, Rui Yang, Quanquan Gu, and Tong Zhang. NeurIPS, 2023.
 - Learning to Modulate pre-trained Models in RL

-
- Thomas Schmied, Markus Hofmarcher, Fabian Paischer, Razvan Pascanu, and Sepp Hochreiter. NeurIPS, 2023.
 - Counterfactual Conservative Q Learning for Offline Multi-agent Reinforcement Learning
 - Jianzhun Shao, Yun Qu, Chen Chen, Hongchang Zhang, and Xiangyang Ji. NeurIPS, 2023.
 - One Risk to Rule Them All: A Risk-Sensitive Perspective on Model-Based Offline Reinforcement Learning
 - Marc Rigter, Bruno Lacerda, and Nick Hawes. NeurIPS, 2023.
 - Goal-Conditioned Predictive Coding for Offline Reinforcement Learning
 - Zilai Zeng, Ce Zhang, Shijie Wang, and Chen Sun. NeurIPS, 2023.
 - Mutual Information Regularized Offline Reinforcement Learning
 - Xiao Ma, Bingyi Kang, Zhongwen Xu, Min Lin, and Shuicheng Yan. NeurIPS, 2023.
 - Offline RL With Heteroskedastic Datasets and Support Constraints
 - Anikait Singh, Aviral Kumar, Quan Vuong, Yevgen Chebotar, and Sergey Levine. NeurIPS, 2023.
 - Offline Reinforcement Learning with Differential Privacy
 - Dan Qiao and Yu-Xiang Wang. NeurIPS, 2023.
 - Accountability in Offline Reinforcement Learning: Explaining Decisions with a Corpus of Examples
 - Hao Sun, Alihan Hüyük, Daniel Jarrett, and Mihaela van der Schaar. NeurIPS, 2023.
 - Reining Generalization in Offline Reinforcement Learning via Representation Distinction
 - Yi Ma, Hongyao Tang, Dong Li, and Zhaopeng Meng. NeurIPS, 2023.
 - VOCE: Variational Optimization with Conservative Estimation for Offline Safe Reinforcement Learning
 - Jiayi Guan, Guang Chen, Jiaming Ji, Long Yang, ao zhou, Zhijun Li, and changjun jiang. NeurIPS, 2023.
 - SafeDICE: Offline Safe Imitation Learning with Non-Preferred Demonstrations
 - Youngsoo Jang, Geon-Hyeong Kim, Jongmin Lee, Sungryull Sohn, Byoungjip Kim, Honglak Lee, and Moontae Lee. NeurIPS, 2023.
 - Hierarchical Diffusion for Offline Decision Making
-

-
- Wenhao Li, Xiangfeng Wang, Bo Jin, and Hongyuan Zha. ICML, 2023.
 - MAHALO: Unifying Offline Reinforcement Learning and Imitation Learning from Observations
 - Anqi Li, Byron Boots, and Ching-An Cheng. ICML, 2023.
 - Safe Offline Reinforcement Learning with Real-Time Budget Constraints
 - Qian Lin, Bo Tang, Zifan Wu, Chao Yu, Shangqin Mao, Qianlong Xie, Xingxing Wang, and Dong Wang. ICML, 2023.
 - Near-optimal Conservative Exploration in Reinforcement Learning under Episode-wise Constraints
 - Donghao Li, Ruiquan Huang, Cong Shen, and Jing Yang. ICML, 2023.
 - A Connection between One-Step Regularization and Critic Regularization in Reinforcement Learning
 - Benjamin Eysenbach, Matthieu Geist, Sergey Levine, and Ruslan Salakhutdinov. ICML, 2023.
 - Anti-Exploration by Random Network Distillation
 - Alexander Nikulin, Vladislav Kurenkov, Denis Tarasov, and Sergey Kolesnikov. ICML, 2023.
 - Optimal Goal-Reaching Reinforcement Learning via Quasimetric Learning
 - Tongzhou Wang, Antonio Torralba, Phillip Isola, and Amy Zhang. ICML, 2023.
 - PASTA: Pessimistic Assortment Optimization
 - Juncheng Dong, Weibin Mo, Zhengling Qi, Cong Shi, Ethan X Fang, and Vahid Tarokh. ICML, 2023.
 - Contrastive Energy Prediction for Exact Energy-Guided Diffusion Sampling in Offline Reinforcement Learning
 - Cheng Lu, Huayu Chen, Jianfei Chen, Hang Su, Chongxuan Li, and Jun Zhu. ICML, 2023.
 - Supported Trust Region Optimization for Offline Reinforcement Learning
 - Yixiu Mao, Hongchang Zhang, Chen Chen, Yi Xu, and Xiangyang Ji. ICML, 2023.
 - Principled Offline RL in the Presence of Rich Exogenous Information
 - Riashat Islam, Manan Tomar, Alex Lamb, Yonathan Efroni, Hongyu Zang, Aniket Rajiv Doldkar, Dipendra Misra, Xin Li, Harm van Seijen, Remi Tachet des Combes, and John Langford. ICML, 2023.
-

-
- Efficient Online Reinforcement Learning with Offline Data
 - Philip J. Ball, Laura Smith, Ilya Kostrikov, and Sergey Levine. ICML, 2023.
 - Boosting Offline Reinforcement Learning with Action Preference Query
 - Qisen Yang, Shenzhi Wang, Matthieu Gaetan Lin, Shiji Song, and Gao Huang. ICML, 2023.
 - Model-based Offline Reinforcement Learning with Count-based Conservatism
 - Byeongchan Kim and Min-hwan Oh. ICML, 2023.
 - Constrained Decision Transformer for Offline Safe Reinforcement Learning
 - Zuxin Liu, Zijian Guo, Yihang Yao, Zhepeng Cen, Wenhao Yu, Tingnan Zhang, and Ding Zhao. ICML, 2023.
 - Model-Bellman Inconsistency for Model-based Offline Reinforcement Learning
 - Yihao Sun, Jiaji Zhang, Chengxing Jia, Haoxin Lin, Junyin Ye, and Yang Yu. ICML, 2023.
 - Provably Efficient Offline Reinforcement Learning with Perturbed Data Sources
 - Chengshuai Shi, Wei Xiong, Cong Shen, and Jing Yang. ICML, 2023.
 - What is Essential for Unseen Goal Generalization of Offline Goal-conditioned RL?
 - Rui Yang, Yong Lin, Xiaoteng Ma, Hao Hu, Chongjie Zhang, and Tong Zhang. ICML, 2023.
 - Policy Regularization with Dataset Constraint for Offline Reinforcement Learning
 - Yuhang Ran, Yi-Chen Li, Fuxiang Zhang, Zongzhang Zhang, and Yang Yu. ICML, 2023.
 - MetaDiffuser: Diffusion Model as Conditional Planner for Offline Meta-RL
 - Fei Ni, Jianye Hao, Yao Mu, Yifu Yuan, Yan Zheng, Bin Wang, and Zhixuan Liang. ICML, 2023.
 - Distance Weighted Supervised Learning for Offline Interaction Data
 - Joey Hejna, Jensen Gao, and Dorsa Sadigh. ICML, 2023.
 - Masked Trajectory Models for Prediction, Representation, and Control
 - Philipp Wu, Arjun Majumdar, Kevin Stone, Yixin Lin, Igor Mordatch, Pieter Abbeel, and Aravind Rajeswaran. ICML, 2023.
 - Contrastive Energy Prediction for Exact Energy-Guided Diffusion Sampling in Offline Reinforcement Learning
 - Cheng Lu, Huayu Chen, Jianfei Chen, Hang Su, Chongxuan Li, and Jun Zhu. ICML, 2023.

-
- Bayesian Reparameterization of Reward-Conditioned Reinforcement Learning with Energy-based Models
 - Wenhao Ding, Tong Che, Ding Zhao, and Marco Pavone. ICML, 2023.
 - Warm-Start Actor-Critic: From Approximation Error to Sub-optimality Gap
 - Hang Wang, Sen Lin, and Junshan Zhang. ICML, 2023.
 - Future-conditioned Unsupervised Pretraining for Decision Transformer
 - Zhihui Xie, Zichuan Lin, Deheng Ye, Qiang Fu, Wei Yang, and Shuai Li. ICML, 2023.
 - PAC-Bayesian Offline Contextual Bandits With Guarantees
 - Otmane Sakhi, Nicolas Chopin, and Pierre Alquier. ICML, 2023.
 - Q-learning Decision Transformer: Leveraging Dynamic Programming for Conditional Sequence Modelling in Offline RL
 - Taku Yamagata, Ahmed Khalil, and Raul Santos-Rodriguez. ICML, 2023.
 - Jump-Start Reinforcement Learning [website]
 - Ikechukwu Uchendu, Ted Xiao, Yao Lu, Banghua Zhu, Mengyuan Yan, Joséphine Simon, Matthew Bennice, Chuyuan Fu, Cong Ma, Jiantao Jiao, Sergey Levine, and Karol Hausman. ICML, 2023.
 - Learning Temporally AbstractWorld Models without Online Experimentation
 - Benjamin Freed, Siddarth Venkatraman, Guillaume Adrien Sartoretti, Jeff Schneider, and Howie Choset. ICML, 2023.
 - A Framework for Adapting Offline Algorithms to Solve Combinatorial Multi-Armed Bandit Problems with Bandit Feedback
 - Guanyu Nie, Yididiya Y Nadew, Yanhui Zhu, Vaneet Aggarwal, and Christopher John Quinn. ICML, 2023.
 - Revisiting the Linear-Programming Framework for Offline RL with General Function Approximation
 - Asuman Ozdaglar, Sarath Pattathil, Jiawei Zhang, and Kaiqing Zhang. ICML, 2023.
 - Semi-Supervised Offline Reinforcement Learning with Action-Free Trajectories
 - Qinqing Zheng, Mikael Henaff, Brandon Amos, and Aditya Grover. ICML, 2023.
 - Actor-Critic Alignment for Offline-to-Online Reinforcement Learning

-
- Zishun Yu and Xinhua Zhang. ICML, 2023.
 - Leveraging Offline Data in Online Reinforcement Learning
 - Andrew Wagenmaker and Aldo Pacchiano. ICML, 2023.
 - Offline Reinforcement Learning with Closed-Form Policy Improvement Operators
 - Jiachen Li, Edwin Zhang, Ming Yin, Qinxun Bai, Yu-Xiang Wang, and William Yang Wang. ICML, 2023.
 - Offline Learning in Markov Games with General Function Approximation
 - Yuheng Zhang, Yu Bai, and Nan Jiang. ICML, 2023.
 - Offline Meta Reinforcement Learning with In-Distribution Online Adaptation
 - Jianhao Wang, Jin Zhang, Haozhe Jiang, Junyu Zhang, Liwei Wang, and Chongjie Zhang. ICML, 2023.
 - Scaling Pareto-Efficient Decision Making Via Offline Multi-Objective RL
 - Baiting Zhu, Meihua Dang, and Aditya Grover. ICLR, 2023.
 - Confidence-Conditioned Value Functions for Offline Reinforcement Learning
 - Joey Hong, Aviral Kumar, and Sergey Levine. ICLR, 2023.
 - Offline Q-Learning on Diverse Multi-Task Data Both Scales And Generalizes [website]
 - Aviral Kumar, Rishabh Agarwal, Xinyang Geng, George Tucker, and Sergey Levine. ICLR, 2023.
 - Is Conditional Generative Modeling all you need for Decision-Making? [website]
 - Anurag Ajay, Yilun Du, Abhi Gupta, Joshua Tenenbaum, Tommi Jaakkola, and Pulkit Agrawal. ICLR, 2023
 - Offline RL with No OOD Actions: In-Sample Learning via Implicit Value Regularization
 - Haoran Xu, Li Jiang, Jianxiong Li, Zhuoran Yang, Zhaoran Wang, Victor Wai Kin Chan, and Xianyu Zhan. ICLR, 2023.
 - Extreme Q-Learning: MaxEnt RL without Entropy
 - Divyansh Garg, Joey Hejna, Matthieu Geist, and Stefano Ermon. ICLR, 2023.
 - Dichotomy of Control: Separating What You Can Control from What You Cannot
 - Mengjiao Yang, Dale Schuurmans, Pieter Abbeel, and Ofir Nachum. ICLR, 2023.
-

-
- From Play to Policy: Conditional Behavior Generation from Uncurated Robot Data
 - Zichen Jeff Cui, Yibin Wang, Nur Muhammad Mahi Shafiullah, and Lerrel Pinto. ICLR, 2023.
 - VIPeR: Provably Efficient Algorithm for Offline RL with Neural Function Approximation
 - Thanh Nguyen-Tang and Raman Arora. ICLR, 2023.
 - Optimal Conservative Offline RL with General Function Approximation via Augmented Lagrangian
 - Paria Rashidinejad, Hanlin Zhu, Kunhe Yang, Stuart Russell, and Jiantao Jiao. ICLR, 2023.
 - The In-Sample Softmax for Offline Reinforcement Learning
 - Chenjun Xiao, Han Wang, Yangchen Pan, Adam White, and Martha White. ICLR, 2023.
 - VIP: Towards Universal Visual Reward and Representation via Value-Implicit Pre-Training [web-site] [code]
 - Yecheng Jason Ma, Shagun Sodhani, Dinesh Jayaraman, Osbert Bastani, Vikash Kumar, and Amy Zhang. ICLR, 2023.
 - Does Zero-Shot Reinforcement Learning Exist?
 - Ahmed Touati, Jérémy Rapin, and Yann Ollivier. ICLR, 2023.
 - Behavior Prior Representation learning for Offline Reinforcement Learning
 - Hongyu Zang, Xin Li, Jie Yu, Chen Liu, Riashat Islam, Remi Tachet Des Combes, and Romain Laroche. ICLR, 2023.
 - Mind the Gap: Offline Policy Optimization for Imperfect Rewards
 - Jianxiong Li, Xiao Hu, Haoran Xu, Jingjing Liu, Xianyuan Zhan, Qing-Shan Jia, and Ya-Qin Zhang. ICLR, 2023.
 - Offline Congestion Games: How Feedback Type Affects Data Coverage Requirement
 - Haozhe Jiang, Qiwen Cui, Zhihan Xiong, Maryam Fazel, and Simon S. Du. ICLR, 2023.
 - User-Interactive Offline Reinforcement Learning
 - Phillip Swazinna, Steffen Udluft, and Thomas Runkler. ICLR, 2023.
 - Discovering Generalizable Multi-agent Coordination Skills from Multi-task Offline Data
 - Fuxiang Zhang, Chengxing Jia, Yi-Chen Li, Lei Yuan, Yang Yu, and Zongzhang Zhang. ICLR, 2023.
-

-
- Hybrid RL: Using Both Offline and Online Data Can Make RL Efficient [code]
 - Yuda Song, Yifei Zhou, Ayush Sekhari, J. Andrew Bagnell, Akshay Krishnamurthy, and Wen Sun. ICLR, 2023.
 - Harnessing Mixed Offline Reinforcement Learning Datasets via Trajectory Weighting
 - Zhang-Wei Hong, Pulkit Agrawal, Remi Tachet des Combes, and Romain Laroche. ICLR, 2023.
 - Efficient Offline Policy Optimization with a Learned Model
 - Zichen Liu, Siyi Li, Wee Sun Lee, Shuicheng Yan, and Zhongwen Xu. ICLR, 2023.
 - Diffusion Policies as an Expressive Policy Class for Offline Reinforcement Learning
 - Zhendong Wang, Jonathan J Hunt, and Mingyuan Zhou. ICLR, 2023.
 - When Data Geometry Meets Deep Function: Generalizing Offline Reinforcement Learning
 - Jianxiong Li, Xianyuan Zhan, Haoran Xu, Xiangyu Zhu, Jingjing Liu, and Ya-Qin Zhang. ICLR, 2023.
 - In-sample Actor Critic for Offline Reinforcement Learning
 - Hongchang Zhang, Yixiu Mao, Boyuan Wang, Shuncheng He, Yi Xu, and Xiangyang Ji. ICLR, 2023.
 - Value Memory Graph: A Graph-Structured World Model for Offline Reinforcement Learning
 - Deyao Zhu, Li Erran Li, and Mohamed Elhoseiny. ICLR, 2023.
 - Conservative Bayesian Model-Based Value Expansion for Offline Policy Optimization
 - Jihwan Jeong, Xiaoyu Wang, Michael Gimelfarb, Hyunwoo Kim, Baher Abdulhai, and Scott Sanner. ICLR, 2023.
 - Offline Reinforcement Learning via High-Fidelity Generative Behavior Modeling
 - Huayu Chen, Cheng Lu, Chengyang Ying, Hang Su, and Jun Zhu. ICLR, 2023.
 - Offline Reinforcement Learning with Differentiable Function Approximation is Provably Efficient
 - Ming Yin, Mengdi Wang, and Yu-Xiang Wang. ICLR, 2023.
 - Nearly Minimax Optimal Offline Reinforcement Learning with Linear Function Approximation: Single-Agent MDP and Markov Game
 - Wei Xiong, Han Zhong, Chengshuai Shi, Cong Shen, Liwei Wang, and Tong Zhang. ICLR, 2023.

-
- Pessimism in the Face of Confounders: Provably Efficient Offline Reinforcement Learning in Partially Observable Markov Decision Processes
 - Miao Lu, Yifei Min, Zhaoran Wang, and Zhuoran Yang. ICLR, 2023.
 - Hyper-Decision Transformer for Efficient Online Policy Adaptation
 - Mengdi Xu, Yuchen Lu, Yikang Shen, Shun Zhang, Ding Zhao, and Chuang Gan. ICLR, 2023.
 - Efficient Planning in a Compact Latent Action Space
 - Zhengyao Jiang, Tianjun Zhang, Michael Janner, Yueying Li, Tim Rocktäschel, Edward Grefenstette, and Yuandong Tian. ICLR, 2023.
 - Preference Transformer: Modeling Human Preferences using Transformers for RL [website]
 - Changyeon Kim, Jongjin Park, Jinwoo Shin, Honglak Lee, Pieter Abbeel, and Kimin Lee. ICLR, 2023.
 - Behavior Proximal Policy Optimization
 - Zifeng Zhuang, Kun Lei, Jinxin Liu, Donglin Wang, and Yilang Guo. ICLR, 2023.
 - Provably Efficient Neural Offline Reinforcement Learning via Perturbed Rewards
 - Thanh Nguyen-Tang and Raman Arora. ICLR, 2023.
 - The Provable Benefits of Unsupervised Data Sharing for Offline Reinforcement Learning
 - Hao Hu, Yiqin Yang, Qianchuan Zhao, and Chongjie Zhang. ICLR, 2023.
 - Decision Transformer under Random Frame Dropping
 - Kaizhe Hu, Ray Chen Zheng, Yang Gao, and Huazhe Xu. ICLR, 2023.
 - Policy Expansion for Bridging Offline-to-Online Reinforcement Learning
 - Haichao Zhang, We Xu, and Haonan Yu. ICLR, 2023.
 - Finetuning Offline World Models in the Real World
 - Yunhai Feng, Nicklas Hansen, Ziyang Xiong, Chandramouli Rajagopalan, and Xiaolong Wang. CoRL, 2023.
 - On the Sample Complexity of Vanilla Model-Based Offline Reinforcement Learning with Dependent Samples
 - Mustafa O. Karabag and Ufuk Topcu. AAAI, 2023.
 - Adaptive Policy Learning for Offline-to-Online Reinforcement Learning
-

-
- Han Zheng, Xufang Luo, Pengfei Wei, Xuan Song, Dongsheng Li, and Jing Jiang. AACL, 2023.
 - Safe Policy Improvement for POMDPs via Finite-State Controllers
 - Thiago D. Simão, Marnix Suilen, and Nils Jansen. AACL, 2023.
 - Behavior Estimation from Multi-Source Data for Offline Reinforcement Learning
 - Guoxi Zhang and Hisashi Kashima. AACL, 2023.
 - On Instance-Dependent Bounds for Offline Reinforcement Learning with Linear Function Approximation
 - Thanh Nguyen-Tang, Ming Yin, Sunil Gupta, Svetha Venkatesh, and Raman Arora. AACL, 2023.
 - Contrastive Example-Based Control
 - Kyle Hatch, Benjamin Eysenbach, Rafael Rafailov, Tianhe Yu, Ruslan Salakhutdinov, Sergey Levine, and Chelsea Finn. LDC, 2023.
 - Curriculum Offline Reinforcement Learning
 - Yuanying Cai, Chuheng Zhang, Hanyue Zhao, Li Zhao, and Jiang Bian. AAMAS. 2023.
 - Offline Reinforcement Learning with On-Policy Q-Function Regularization
 - Laixi Shi, Robert Dadashi, Yuejie Chi, Pablo Samuel Castro, and Matthieu Geist. ECML, 2023.
 - Model-based Offline Policy Optimization with Adversarial Network
 - Junming Yang, Xingguo Chen, Shengyuan Wang, and Bolei Zhang. ECAI, 2023.
 - Efficient experience replay architecture for offline reinforcement learning
 - Longfei Zhang, Yanghe Feng, Rongxiao Wang, Yue Xu, Naifu Xu, Zeyi Liu, and Hang Du. RIA, 2023.
 - Automatic Trade-off Adaptation in Offline RL
 - Phillip Swazinna, Steffen Udluft, and Thomas Runkler. ESANN, 2023.
 - Offline Robot Reinforcement Learning with Uncertainty-Guided Human Expert Sampling
 - Ashish Kumar and Ilya Kuzovkin. arXiv, 2022.
 - Latent Variable Representation for Reinforcement Learning
 - Tongzheng Ren, Chenjun Xiao, Tianjun Zhang, Na Li, Zhaoran Wang, Sujay Sanghavi, Dale Schuurmans, and Bo Dai. arXiv, 2022.
-

-
- Learning From Good Trajectories in Offline Multi-Agent Reinforcement Learning
 - Qi Tian, Kun Kuang, Furui Liu, and Baoxiang Wang. arXiv, 2022.
 - State-Aware Proximal Pessimistic Algorithms for Offline Reinforcement Learning
 - Chen Chen, Hongyao Tang, Yi Ma, Chao Wang, Qianli Shen, Dong Li, and Jianye Hao. arXiv, 2022.
 - Masked Autoencoding for Scalable and Generalizable Decision Making
 - Fangchen Liu, Hao Liu, Aditya Grover, and Pieter Abbeel. arXiv, 2022.
 - Improving TD3-BC: Relaxed Policy Constraint for Offline Learning and Stable Online Fine-Tuning
 - Alex Beeson and Giovanni Montana. arXiv, 2022.
 - Q-Ensemble for Offline RL: Don’t Scale the Ensemble, Scale the Batch Size
 - Alexander Nikulin, Vladislav Kurenkov, Denis Tarasov, Dmitry Akimov, and Sergey Kolesnikov. arXiv, 2022.
 - Let Offline RL Flow: Training Conservative Agents in the Latent Space of Normalizing Flows
 - Dmitriy Akimov, Vladislav Kurenkov, Alexander Nikulin, Denis Tarasov, and Sergey Kolesnikov. arXiv, 2022.
 - Model-based Trajectory Stitching for Improved Offline Reinforcement Learning
 - Charles A. Hepburn and Giovanni Montana. arXiv, 2022.
 - Offline Reinforcement Learning with Adaptive Behavior Regularization
 - Yunfan Zhou, Xijun Li, and Qingyu Qu. arXiv, 2022.
 - Contextual Transformer for Offline Meta Reinforcement Learning
 - Runji Lin, Ye Li, Xidong Feng, Zhaowei Zhang, Xian Hong Wu Fung, Haifeng Zhang, Jun Wang, Yali Du, and Yaodong Yang. arXiv, 2022.
 - Wall Street Tree Search: Risk-Aware Planning for Offline Reinforcement Learning
 - Dan Elbaz, Gal Novik, and Oren Salzman. arXiv, 2022.
 - ARMOR: A Model-based Framework for Improving Arbitrary Baseline Policies with Offline Data
 - Tengyang Xie, Mohak Bhardwaj, Nan Jiang, and Ching-An Cheng. arXiv, 2022.
 - Contrastive Value Learning: Implicit Models for Simple Offline RL

-
- Bogdan Mazouze, Benjamin Eysenbach, Ofir Nachum, and Jonathan Tompson. arXiv, 2022.
 - Optimistic Curiosity Exploration and Conservative Exploitation with Linear Reward Shaping
 - Hao Sun, Lei Han, Rui Yang, Xiaoteng Ma, Jian Guo, and Bolei Zhou. arXiv, 2022.
 - Optimal Conservative Offline RL with General Function Approximation via Augmented Lagrangian
 - Paria Rashidinejad, Hanlin Zhu, Kunhe Yang, Stuart Russell, and Jiantao Jiao. ICLR, 2023.
 - Agent-Controller Representations: Principled Offline RL with Rich Exogenous Information
 - Riashat Islam, Manan Tomar, Alex Lamb, Yonathan Efroni, Hongyu Zang, Aniket Didolkar, Dipendra Misra, Xin Li, Harm van Seijen, Remi Tachet des Combes, and John Langford. arXiv, 2022.
 - Provable Safe Reinforcement Learning with Binary Feedback
 - Andrew Bennett, Dipendra Misra, and Nathan Kallus. arXiv, 2022.
 - Learning on the Job: Self-Rewarding Offline-to-Online Finetuning for Industrial Insertion of Novel Connectors from Vision
 - Ashvin Nair, Brian Zhu, Gokul Narayanan, Eugen Solowjow, and Sergey Levine. arXiv, 2022.
 - Implicit Offline Reinforcement Learning via Supervised Learning
 - Alexandre Piche, Rafael Pardinias, David Vazquez, Igor Mordatch, and Chris Pal. arXiv, 2022.
 - Robust Offline Reinforcement Learning with Gradient Penalty and Constraint Relaxation
 - Chengqian Gao, Ke Xu, Liu Liu, Deheng Ye, Peilin Zhao, and Zhiqiang Xu. arXiv, 2022.
 - Boosting Offline Reinforcement Learning via Data Rebalancing
 - Yang Yue, Bingyi Kang, Xiao Ma, Zhongwen Xu, Gao Huang, and Shuicheng Yan. arXiv, 2022.
 - ConserWeightive Behavioral Cloning for Reliable Offline Reinforcement Learning [code]
 - Tung Nguyen, Qinqing Zheng, and Aditya Grover. arXiv, 2022.
 - State Advantage Weighting for Offline RL
 - Jiafei Lyu, Aicheng Gong, Le Wan, Zongqing Lu, and Xiu Li. arXiv, 2022.
 - Blessing from Experts: Super Reinforcement Learning in Confounded Environments
 - Jiayi Wang, Zhengling Qi, and Chengchun Shi. arXiv, 2022.
-

-
- DCE: Offline Reinforcement Learning With Double Conservative Estimates
 - Chen Zhao, Kai Xing Huang, and Chun Yuan. arXiv, 2022.
 - On the Opportunities and Challenges of using Animals Videos in Reinforcement Learning
 - Vittorio Giammarino. arXiv, 2022.
 - Offline Reinforcement Learning with Instrumental Variables in Confounded Markov Decision Processes
 - Zuyue Fu, Zhengling Qi, Zhaoran Wang, Zhuoran Yang, Yanxun Xu, and Michael R. Kosorok. arXiv, 2022.
 - Exploiting Reward Shifting in Value-Based Deep RL
 - Hao Sun, Lei Han, Rui Yang, Xiaoteng Ma, Jian Guo, and Bolei Zhou. arXiv, 2022.
 - Distributionally Robust Offline Reinforcement Learning with Linear Function Approximation
 - Xiaoteng Ma, Zhipeng Liang, Li Xia, Jiheng Zhang, Jose Blanchet, Mingwen Liu, Qianchuan Zhao, and Zhengyuan Zhou. arXiv, 2022.
 - C²: Co-design of Robots via Concurrent Networks Coupling Online and Offline Reinforcement Learning
 - Ci Chen, Pingyu Xiang, Haojian Lu, Yue Wang, and Rong Xiong. arXiv, 2022.
 - Strategic Decision-Making in the Presence of Information Asymmetry: Provably Efficient RL with Algorithmic Instruments
 - Mengxin Yu, Zhuoran Yang, and Jianqing Fan. arXiv, 2022.
 - Distributionally Robust Model-Based Offline Reinforcement Learning with Near-Optimal Sample Complexity
 - Laixi Shi and Yuejie Chi. arXiv, 2022.
 - AdaCat: Adaptive Categorical Discretization for Autoregressive Models
 - Qiyang Li, Ajay Jain, and Pieter Abbeel. arXiv, 2022.
 - Branch Ranking for Efficient Mixed-Integer Programming via Offline Ranking-based Policy Learning
 - Zeren Huang, Wenhao Chen, Weinan Zhang, Chuhan Shi, Furui Liu, Hui-Ling Zhen, Mingxuan Yuan, Jianye Hao, Yong Yu, and Jun Wang. arXiv, 2022.
 - Offline Reinforcement Learning at Multiple Frequencies [webpage]
-

-
- Kaylee Burns, Tianhe Yu, Chelsea Finn, and Karol Hausman. arXiv, 2022.
 - General Policy Evaluation and Improvement by Learning to Identify Few But Crucial States
 - Francesco Faccio, Aditya Ramesh, Vincent Herrmann, Jean Harb, and Jürgen Schmidhuber. arXiv, 2022.
 - Behavior Transformers: Cloning k modes with one stone
 - Nur Muhammad Mahi Shafiullah, Zichen Jeff Cui, Ariuntuya Altanzaya, and Lerrel Pinto. arXiv, 2022.
 - Contrastive Learning as Goal-Conditioned Reinforcement Learning
 - Benjamin Eysenbach, Tianjun Zhang, Ruslan Salakhutdinov, and Sergey Levine. arXiv, 2022.
 - Federated Offline Reinforcement Learning
 - Doudou Zhou, Yufeng Zhang, Aaron Sonabend-W, Zhaoran Wang, Junwei Lu, and Tianxi Cai. arXiv, 2022.
 - Provable Benefit of Multitask Representation Learning in Reinforcement Learning
 - Yuan Cheng, Songtao Feng, Jing Yang, Hong Zhang, and Yingbin Liang. arXiv, 2022.
 - Provably Efficient Offline Reinforcement Learning with Trajectory-Wise Reward
 - Tengyu Xu and Yingbin Liang. arXiv, 2022.
 - Model-Based Reinforcement Learning Is Minimax-Optimal for Offline Zero-Sum Markov Games
 - Yuling Yan, Gen Li, Yuxin Chen, and Jianqing Fan. arXiv, 2022.
 - Offline Reinforcement Learning with Causal Structured World Models
 - Zheng-Mao Zhu, Xiong-Hui Chen, Hong-Long Tian, Kun Zhang, and Yang Yu. arXiv, 2022.
 - Incorporating Explicit Uncertainty Estimates into Deep Offline Reinforcement Learning
 - David Brandfonbrener, Remi Tachet des Combes, and Romain Laroche. arXiv, 2022.
 - Know Your Boundaries: The Necessity of Explicit Behavioral Cloning in Offline RL
 - Wonjoon Goo and Scott Niekum. arXiv, 2022.
 - Byzantine-Robust Online and Offline Distributed Reinforcement Learning
 - Yiding Chen, Xuezhou Zhang, Kaiqing Zhang, Mengdi Wang, and Xiaojin Zhu. arXiv, 2022.
 - Model Generation with Provable Coverability for Offline Reinforcement Learning
-

-
- Chengxing Jia, Hao Yin, Chenxiao Gao, Tian Xu, Lei Yuan, Zongzhang Zhang, and Yang Yu. arXiv, 2022.
 - You Can’t Count on Luck: Why Decision Transformers Fail in Stochastic Environments
 - Keiran Paster, Sheila McIlraith, and Jimmy Ba. arXiv, 2022.
 - Multi-Game Decision Transformers
 - Kuang-Huei Lee, Ofir Nachum, Mengjiao Yang, Lisa Lee, Daniel Freeman, Winnie Xu, Sergio Guadarrama, Ian Fischer, Eric Jang, Henryk Michalewski, and Igor Mordatch. arXiv, 2022.
 - Hierarchical Planning Through Goal-Conditioned Offline Reinforcement Learning
 - Jinning Li, Chen Tang, Masayoshi Tomizuka, and Wei Zhan. arXiv, 2022.
 - Distance-Sensitive Offline Reinforcement Learning
 - Jianxiong Li, Xianyuan Zhan, Haoran Xu, Xiangyu Zhu, Jingjing Liu, and Ya-Qin Zhang. arXiv, 2022.
 - No More Pesky Hyperparameters: Offline Hyperparameter Tuning for RL
 - Han Wang, Archit Saxhadeo, Adam White, James Bell, Vincent Liu, Xutong Zhao, Puer Liu, Tadashi Kozuno, Alona Fyshe, and Martha White. arXiv, 2022.
 - How to Spend Your Robot Time: Bridging Kickstarting and Offline Reinforcement Learning for Vision-based Robotic Manipulation
 - Alex X. Lee, Coline Devin, Jost Tobias Springenberg, Yuxiang Zhou, Thomas Lampe, Abbas Abdolmaleki, and Konstantinos Bousmalis. arXiv, 2022.
 - Offline Visual Representation Learning for Embodied Navigation
 - Karmesh Yadav, Ram Ramrakhya, Arjun Majumdar, Vincent-Pierre Berges, Sachit Kuhar, Dhruv Batra, Alexei Baevski, and Oleksandr Maksymets. arXiv, 2022.
 - Towards Flexible Inference in Sequential Decision Problems via Bidirectional Transformers
 - Micah Carroll, Jessy Lin, Orr Paradise, Raluca Georgescu, Mingfei Sun, David Bignell, Stephanie Milani, Katja Hofmann, Matthew Hausknecht, Anca Dragan, and Sam Devlin. arXiv, 2022.
 - BATS: Best Action Trajectory Stitching
 - Ian Char, Viraj Mehta, Adam Villafior, John M. Dolan, Jeff Schneider. arXiv, 2022.
 - Settling the Sample Complexity of Model-Based Offline Reinforcement Learning
-

-
- Gen Li, Laixi Shi, Yuxin Chen, Yuejie Chi, and Yuting Wei. arXiv, 2022.
 - PAnDR: Fast Adaptation to New Environments from Offline Experiences via Decoupling Policy and Environment Representations
 - Tong Sang, Hongyao Tang, Yi Ma, Jianye Hao, Yan Zheng, Zhaopeng Meng, Boyan Li, and Zhen Wang. arXiv, 2022.
 - Offline Reinforcement Learning Under Value and Density-Ratio Realizability: the Power of Gaps
 - Jinglin Chen and Nan Jiang. arXiv, 2022.
 - Meta Reinforcement Learning for Adaptive Control: An Offline Approach
 - Daniel G. McClement, Nathan P. Lawrence, Johan U. Backstrom, Philip D. Loewen, Michael G. Forbes, and R. Bhushan Gopaluni. arXiv, 2022.
 - The Efficacy of Pessimism in Asynchronous Q-Learning
 - Yuling Yan, Gen Li, Yuxin Chen, and Jianqing Fan. arXiv, 2022.
 - Reinforcement Learning for Linear Quadratic Control is Vulnerable Under Cost Manipulation
 - Yunhan Huang and Quanyan Zhu. arXiv, 2022.
 - A Regularized Implicit Policy for Offline Reinforcement Learning
 - Shentao Yang, Zhendong Wang, Huangjie Zheng, Yihao Feng, and Mingyuan Zhou. arXiv, 2022.
 - Reinforcement Learning in Possibly Nonstationary Environments [code]
 - Mengbing Li, Chengchun Shi, Zhenke Wu, and Piotr Fryzlewicz. arXiv, 2022.
 - Statistically Efficient Advantage Learning for Offline Reinforcement Learning in Infinite Horizons
 - Chengchun Shi, Shikai Luo, Hongtu Zhu, and Rui Song. arXiv, 2022.
 - VRL3: A Data-Driven Framework for Visual Deep Reinforcement Learning
 - Che Wang, Xufang Luo, Keith Ross, and Dongsheng Li. arXiv, 2022.
 - Retrieval-Augmented Reinforcement Learning
 - Anirudh Goyal, Abram L. Friesen, Andrea Banino, Theophane Weber, Nan Rosemary Ke, Adria Puigdomenech Badia, Arthur Guez, Mehdi Mirza, Ksenia Konyushkova, Michal Valko, Simon Osindero, Timothy Lillicrap, Nicolas Heess, and Charles Blundell. arXiv, 2022.
 - Online Decision Transformer

-
- Qinqing Zheng, Amy Zhang, and Aditya Grover. arXiv, 2022.
 - Transferred Q-learning
 - Elynn Y. Chen, Michael I. Jordan, and Sai Li. arXiv, 2022.
 - Settling the Communication Complexity for Distributed Offline Reinforcement Learning
 - Juliusz Krysztof Ziomek, Jun Wang, and Yaodong Yang. arXiv, 2022.
 - Offline Reinforcement Learning with Realizability and Single-policy Concentrability
 - Wenhao Zhan, Baihe Huang, Audrey Huang, Nan Jiang, and Jason D. Lee. arXiv, 2022.
 - Rethinking Goal-conditioned Supervised Learning and Its Connection to Offline RL
 - Rui Yang, Yiming Lu, Wenzhe Li, Hao Sun, Meng Fang, Yali Du, Xiu Li, Lei Han, and Chongjie Zhang. arXiv, 2022.
 - Stochastic Gradient Descent with Dependent Data for Offline Reinforcement Learning
 - Jing Dong and Xin T. Tong. arXiv, 2022.
 - Can Wikipedia Help Offline Reinforcement Learning?
 - Machel Reid, Yutaro Yamada, and Shixiang Shane Gu. arXiv, 2022.
 - MOORe: Model-based Offline-to-Online Reinforcement Learning
 - Yihuan Mao, Chao Wang, Bin Wang, and Chongjie Zhang. arXiv, 2022.
 - Operator Deep Q-Learning: Zero-Shot Reward Transferring in Reinforcement Learning
 - Ziyang Tang, Yihao Feng, and Qiang Liu. arXiv, 2022.
 - Importance of Empirical Sample Complexity Analysis for Offline Reinforcement Learning
 - Samin Yeasar Arnob, Riashat Islam, and Doina Precup. arXiv, 2022.
 - Single-Shot Pruning for Offline Reinforcement Learning
 - Samin Yeasar Arnob, Riyasat Ohib, Sergey Plis, and Doina Precup. arXiv, 2022.
 - Monte Carlo Augmented Actor-Critic for Sparse Reward Deep Reinforcement Learning from Sub-optimal Demonstrations [website] [code]
 - Albert Wilcox, Ashwin Balakrishna, Jules Dedieu, Wyame Benslimane, Daniel S. Brown, and Ken Goldberg. NeurIPS, 2022.
 - Data-Driven Offline Decision-Making via Invariant Representation Learning

-
- Han Qi, Yi Su, Aviral Kumar, and Sergey Levine. NeurIPS, 2022.
 - Bellman Residual Orthogonalization for Offline Reinforcement Learning
 - Andrea Zanette, and Martin J. Wainwright. NeurIPS, 2022.
 - A Near-Optimal Primal-Dual Method for Off-Policy Learning in CMDP
 - Fan Chen, Junyu Zhang, and Zaiwen Wen. NeurIPS, 2022.
 - RORL: Robust Offline Reinforcement Learning via Conservative Smoothing
 - Rui Yang, Chenjia Bai, Xiaoteng Ma, Zhaoran Wang, Chongjie Zhang, and Lei Han. NeurIPS, 2022.
 - On Gap-dependent Bounds for Offline Reinforcement Learning
 - Xinqi Wang, Qiwen Cui, and Simon S. Du. NeurIPS, 2022.
 - Provably Efficient Offline Multi-agent Reinforcement Learning via Strategy-wise Bonus
 - Qiwen Cui and Simon S. Du. NeurIPS, 2022.
 - Supported Policy Optimization for Offline Reinforcement Learning
 - Jialong Wu, Haixu Wu, Zihan Qiu, Jianmin Wang, and Mingsheng Long. NeurIPS, 2022.
 - When to Trust Your Simulator: Dynamics-Aware Hybrid Offline-and-Online Reinforcement Learning
 - Haoyi Niu, Shubham Sharma, Yiwen Qiu, Ming Li, Guyue Zhou, Jianming Hu, and Xianyuan Zhan. NeurIPS, 2022.
 - Why So Pessimistic? Estimating Uncertainties for Offline RL through Ensembles, and Why Their Independence Matters
 - Seyed Kamyar Seyed Ghasemipour, Shixiang Shane Gu, and Ofir Nachum. NeurIPS, 2022.
 - When does return-conditioned supervised learning work for offline reinforcement learning?
 - David Brandfonbrener, Alberto Bietti, Jacob Buckman, Romain Laroche, and Joan Bruna. NeurIPS, 2022.
 - Pessimism for Offline Linear Contextual Bandits using ℓ_p Confidence Sets
 - Gene Li, Cong Ma, and Nathan Srebro. NeurIPS, 2022.
 - RAMBO-RL: Robust Adversarial Model-Based Offline Reinforcement Learning
 - Marc Rigter, Bruno Lacerda, and Nick Hawes. NeurIPS, 2022.

-
- When is Offline Two-Player Zero-Sum Markov Game Solvable?
 - Qiwen Cui, and Simon S. Du. NeurIPS, 2022.
 - Robust Reinforcement Learning using Offline Data
 - Kishan Panaganti, Zaiyan Xu, Dileep Kalathil, and Mohammad Ghavamzadeh. NeurIPS, 2022.
 - Bidirectional Learning for Offline Infinite-width Model-based Optimization
 - Can Chen, Yingxue Zhang, Jie Fu, Xue Liu, and Mark Coates. NeurIPS, 2022.
 - Mildly Conservative Q-Learning for Offline Reinforcement Learning
 - Jiafei Lyu, Xiaoteng Ma, Xiu Li, and Zongqing Lu. NeurIPS, 2022.
 - Bootstrapped Transformer for Offline Reinforcement Learning
 - Kerong Wang, Hanye Zhao, Xufang Luo, Kan Ren, Weinan Zhang, and Dongsheng Li. NeurIPS, 2022.
 - LobsDICE: Offline Learning from Observation via Stationary Distribution Correction Estimation
 - Geon-Hyeong Kim, Jongmin Lee, Youngsoo Jang, Hongseok Yang, and Kee-Eung Kim. NeurIPS, 2022.
 - Latent-Variable Advantage-Weighted Policy Optimization for Offline RL
 - Xi Chen, Ali Ghadirzadeh, Tianhe Yu, Yuan Gao, Jianhao Wang, Wenzhe Li, Bin Liang, Chelsea Finn, and Chongjie Zhang. NeurIPS, 2022.
 - Double Check Your State Before Trusting It: Confidence-Aware Bidirectional Offline Model-Based Imagination
 - Jiafei Lyu, Xiu Li, and Zongqing Lu. NeurIPS, 2022.
 - Improving Zero-shot Generalization in Offline Reinforcement Learning using Generalized Similarity Functions
 - Bogdan Mazouze, Ilya Kostrikov, Ofir Nachum, and Jonathan Tompson. NeurIPS, 2022.
 - Offline Goal-Conditioned Reinforcement Learning via f-Advantage Regression
 - Yecheng Jason Ma, Jason Yan, Dinesh Jayaraman, and Osbert Bastani. NeurIPS, 2022.
 - Dual Generator Offline Reinforcement Learning
 - Quan Vuong, Aviral Kumar, Sergey Levine, and Yevgen Chebotar. NeurIPS, 2022.

-
- MoCoDA: Model-based Counterfactual Data Augmentation
 - Silviu Pitis, Elliot Creager, Ajay Mandlekar, and Animesh Garg. NeurIPS, 2022.
 - A Policy-Guided Imitation Approach for Offline Reinforcement Learning [code]
 - Haoran Xu, Li Jiang, Jianxiong Li, and Xianyuan Zhan. NeurIPS, 2022.
 - A Unified Framework for Alternating Offline Model Training and Policy Learning
 - Shentao Yang, Shujian Zhang, Yihao Feng, and Mingyuan Zhou. NeurIPS, 2022.
 - Model-Based Offline Reinforcement Learning with Pessimism-Modulated Dynamics Belief
 - Kaiyang Guo, Yunfeng Shao, and Yanhui Geng. NeurIPS, 2022.
 - S2P: State-conditioned Image Synthesis for Data Augmentation in Offline Reinforcement Learning
 - Daesol Cho, Dongseok Shim, and H. Jin Kim. NeurIPS, 2022.
 - ASPIRe: Adaptive Skill Priors for Reinforcement Learning
 - Mengda Xu, Manuela Veloso, and Shuran Song. NeurIPS, 2022.
 - Skills Regularized Task Decomposition for Multi-task Offline Reinforcement Learning
 - Minjong Yoo, Sangwoo Cho, and Honguk Woo. NeurIPS, 2022.
 - Offline Multi-Agent Reinforcement Learning with Knowledge Distillation
 - Wei-Cheng Tseng, Tsun-Hsuan Wang, Yen-Chen Lin, and Phillip Isola. NeurIPS, 2022.
 - Shadow Knowledge Distillation: Bridging Offline and Online Knowledge Transfer
 - Lujun Li and Zhe Jin. NeurIPS, 2022.
 - Addressing Optimism Bias in Sequence Modeling for Reinforcement Learning
 - Adam Villaflor, Zhe Huang, Swapnil Pande, John Dolan, and Jeff Schneider. ICML, 2022.
 - Offline RL Policies Should be Trained to be Adaptive
 - Dibya Ghosh, Anurag Ajay, Pulkit Agrawal, and Sergey Levine. ICML, 2022.
 - Adversarially Trained Actor Critic for Offline Reinforcement Learning
 - Ching-An Cheng, Tengyang Xie, Nan Jiang, and Alekh Agarwal. ICML, 2022.
 - Pessimistic Minimax Value Iteration: Provably Efficient Equilibrium Learning from Offline Datasets

-
- Han Zhong, Wei Xiong, Jiyuan Tan, Liwei Wang, Tong Zhang, Zhaoran Wang, and Zhuoran Yang. ICML, 2022.
 - How to Leverage Unlabeled Data in Offline Reinforcement Learning
 - Tianhe Yu, Aviral Kumar, Yevgen Chebotar, Karol Hausman, Chelsea Finn, and Sergey Levine. ICML, 2022.
 - Plan Better Amid Conservatism: Offline Multi-Agent Reinforcement Learning with Actor Rectification
 - Ling Pan, Longbo Huang, Tengyu Ma, and Huazhe Xu. ICML, 2022.
 - Learning Pseudometric-based Action Representations for Offline Reinforcement Learning
 - Pengjie Gu, Mengchen Zhao, Chen Chen, Dong Li, Jianye Hao, and Bo An. ICML, 2022.
 - Offline Meta-Reinforcement Learning with Online Self-Supervision
 - Vitchyr H. Pong, Ashvin Nair, Laura Smith, Catherine Huang, and Sergey Levine. ICML, 2022.
 - Versatile Offline Imitation from Observations and Examples via Regularized State-Occupancy Matching
 - Yecheng Jason Ma, Andrew Shen, Dinesh Jayaraman, and Osbert Bastani. ICML, 2022.
 - Constrained Offline Policy Optimization
 - Nicholas Polosky, Bruno C. Da Silva, Madalina Fiterau, and Jithin Jagannath. ICML, 2022.
 - Discriminator-Weighted Offline Imitation Learning from Suboptimal Demonstrations
 - Haoran Xu, Xianyuan Zhan, Honglei Yin, and Huiling Qin. ICML, 2022.
 - Provably Efficient Offline Reinforcement Learning for Partially Observable Markov Decision Processes
 - Hongyi Guo, Qi Cai, Yufeng Zhang, Zhuoran Yang, and Zhaoran Wang. ICML, 2022.
 - Pessimistic Q-Learning for Offline Reinforcement Learning: Towards Optimal Sample Complexity
 - Laixi Shi, Gen Li, Yuting Wei, Yuxin Chen, and Yuejie Chi. ICML, 2022.
 - Efficient Reinforcement Learning in Block MDPs: A Model-free Representation Learning Approach
 - Xuezhou Zhang, Yuda Song, Masatoshi Uehara, Mengdi Wang, Alekh Agarwal, and Wen Sun. ICML, 2022.
-

-
- Prompting Decision Transformer for Few-Shot Policy Generalization
 - Mengdi Xu, Yikang Shen, Shun Zhang, Yuchen Lu, Ding Zhao, Joshua B. Tenenbaum, and Chuang Gan. ICML, 2022.
 - Regularizing a Model-based Policy Stationary Distribution to Stabilize Offline Reinforcement Learning
 - Shentao Yang, Yihao Feng, Shujian Zhang, and Mingyuan Zhou. ICML, 2022.
 - On the Role of Discount Factor in Offline Reinforcement Learning
 - Hao Hu, Yiqin Yang, Qianchuan Zhao, and Chongjie Zhang. ICML, 2022.
 - Koopman Q-learning: Offline Reinforcement Learning via Symmetries of Dynamics
 - Matthias Weissenbacher, Samarth Sinha, Animesh Garg, and Yoshinobu Kawahara. ICML, 2022.
 - Representation Learning for Online and Offline RL in Low-rank MDPs [video]
 - Masatoshi Uehara, Xuezhou Zhang, and Wen Sun. ICLR, 2022.
 - Pessimistic Model-based Offline Reinforcement Learning under Partial Coverage [video]
 - Masatoshi Uehara and Wen Sun. ICLR, 2022.
 - Revisiting Design Choices in Model-Based Offline Reinforcement Learning
 - Cong Lu, Philip J. Ball, Jack Parker-Holder, Michael A. Osborne, and Stephen J. Roberts. ICLR, 2022.
 - DR3: Value-Based Deep Reinforcement Learning Requires Explicit Regularization
 - Aviral Kumar, Rishabh Agarwal, Tengyu Ma, Aaron Courville, George Tucker, and Sergey Levine. ICLR, 2022.
 - COptiDICE: Offline Constrained Reinforcement Learning via Stationary Distribution Correction Estimation
 - Jongmin Lee, Cosmin Paduraru, Daniel J. Mankowitz, Nicolas Heess, Doina Precup, Kee-Eung Kim, and Arthur Guez. ICLR, 2022.
 - POETREE: Interpretable Policy Learning with Adaptive Decision Trees
 - Alizée Pace, Alex J. Chan, and Mihaela van der Schaar. ICLR, 2022.
 - Planning in Stochastic Environments with a Learned Model

-
- Ioannis Antonoglou, Julian Schrittwieser, Sherjil Ozair, Thomas K Hubert, and David Silver. ICLR, 2022.
 - Offline Reinforcement Learning with Value-based Episodic Memory
 - Xiaoteng Ma, Yiqin Yang, Hao Hu, Qihan Liu, Jun Yang, Chongjie Zhang, Qianchuan Zhao, and Bin Liang. ICLR, 2022.
 - When Should We Prefer Offline Reinforcement Learning Over Behavioral Cloning?
 - Aviral Kumar, Joey Hong, Anikait Singh, and Sergey Levine. ICLR, 2022.
 - Learning Value Functions from Undirected State-only Experience [website] [code]
 - Matthew Chang, Arjun Gupta, and Saurabh Gupta. ICLR, 2022.
 - Rethinking Goal-Conditioned Supervised Learning and Its Connection to Offline RL
 - Rui Yang, Yiming Lu, Wenzhe Li, Hao Sun, Meng Fang, Yali Du, Xiu Li, Lei Han, and Chongjie Zhang. ICLR, 2022.
 - Offline Reinforcement Learning with Implicit Q-Learning
 - Ilya Kostrikov, Ashvin Nair, and Sergey Levine. ICLR, 2022.
 - RvS: What is Essential for Offline RL via Supervised Learning?
 - Scott Emmons, Benjamin Eysenbach, Ilya Kostrikov, and Sergey Levine. ICLR, 2022.
 - Pareto Policy Pool for Model-based Offline Reinforcement Learning
 - Yijun Yang, Jing Jiang, Tianyi Zhou, Jie Ma, and Yuhui Shi. ICLR, 2022.
 - CrowdPlay: Crowdsourcing Human Demonstrations for Offline Learning
 - Matthias Gerstgrasser, Rakshit Trivedi, and David C. Parkes. ICLR, 2022.
 - COPA: Certifying Robust Policies for Offline Reinforcement Learning against Poisoning Attacks
 - Fan Wu, Linyi Li, Chejian Xu, Huan Zhang, Bhavya Kailkhura, Krishnaram Kenthapadi, Ding Zhao, and Bo Li. ICLR, 2022.
 - DARA: Dynamics-Aware Reward Augmentation in Offline Reinforcement Learning
 - Jinxin Liu, Hongyin Zhang, and Donglin Wang. ICLR, 2022.
 - Near-optimal Offline Reinforcement Learning with Linear Representation: Leveraging Variance Information with Pessimism
 - Ming Yin, Yaqi Duan, Mengdi Wang, and Yu-Xiang Wang. ICLR, 2022.
-

-
- Pessimistic Bootstrapping for Uncertainty-Driven Offline Reinforcement Learning
 - Chenjia Bai, Lingxiao Wang, Zhuoran Yang, Zhihong Deng, Animesh Garg, Peng Liu, and Zhaoran Wang. ICLR, 2022.
 - Offline Neural Contextual Bandits: Pessimism, Optimization and Generalization
 - Thanh Nguyen-Tang, Sunil Gupta, A.Tuan Nguyen, and Svetha Venkatesh. ICLR, 2022.
 - Generalized Decision Transformer for Offline Hindsight Information Matching [website]
 - Hiroki Furuta, Yutaka Matsuo, and Shixiang Shane Gu. ICLR, 2022.
 - Model-Based Offline Meta-Reinforcement Learning with Regularization
 - Sen Lin, Jialin Wan, Tengyu Xu, Yingbin Liang, and Junshan Zhang. ICLR, 2022.
 - AW-Opt: Learning Robotic Skills with Imitation and Reinforcement at Scale [website]
 - Yao Lu, Karol Hausman, Yevgen Chebotar, Mengyuan Yan, Eric Jang, Alexander Herzog, Ted Xiao, Alex Irpan, Mohi Khansari, Dmitry Kalashnikov, and Sergey Levine. CoRL, 2022.
 - Dealing with the Unknown: Pessimistic Offline Reinforcement Learning
 - Jinning Li, Chen Tang, Masayoshi Tomizuka, and Wei Zhan. CoRL, 2022.
 - You Only Evaluate Once: a Simple Baseline Algorithm for Offline RL
 - Wonjoon Goo and Scott Niekum. CoRL, 2022.
 - S4RL: Surprisingly Simple Self-Supervision for Offline Reinforcement Learning
 - Samarth Sinha and Animesh Garg. CoRL, 2022.
 - A Workflow for Offline Model-Free Robotic Reinforcement Learning [website]
 - Aviral Kumar, Anikait Singh, Stephen Tian, Chelsea Finn, and Sergey Levine. CoRL, 2022.
 - Beyond Pick-and-Place: Tackling Robotic Stacking of Diverse Shapes [blog] [video] [code]
 - Alex X. Lee, Coline Devin, Yuxiang Zhou, Thomas Lampe, Konstantinos Bousmalis, Jost Tobias Springenberg, Arunkumar Byravan, Abbas Abdolmaleki, Nimrod Gileadi, David Khosid, Claudio Fantacci, Jose Enrique Chen, Akhil Raju, Rae Jeong, Michael Neunert, Antoine Laurens, Stefano Saliceti, Federico Casarini, Martin Riedmiller, Raia Hadsell, and Francesco Nori. CoRL, 2022.
 - Finetuning from Offline Reinforcement Learning: Challenges, Trade-offs and Practical Solutions
 - Yicheng Luo, Jackie Kay, Edward Grefenstette, and Marc Peter Deisenroth. RLDM, 2022.

-
- Offline Reinforcement Learning with Representations for Actions
 - Xingzhou Lou, Qiyue Yin, Junge Zhang, Chao Yu, Zhaofeng He, Nengjie Cheng, and Kaiqi Huang. Information Sciences, 2022.
 - Towards Off-Policy Learning for Ranking Policies with Logged Feedback
 - Teng Xiao and Suhan Wang. AAAI, 2022.
 - Safe Offline Reinforcement Learning Through Hierarchical Policies
 - Shaofan Liu and Shiliang Sun. PAKDD, 2022.
 - TD3 with Reverse KL Regularizer for Offline Reinforcement Learning from Mixed Datasets
 - Yuanying Cai, Chuheng Zhang, Li Zhao, Wei Shen, Xuyun Zhang, Lei Song, Jiang Bian, Tao Qin, and Tieyan Liu. ICDM, 2022.
 - Sample Complexity of Offline Reinforcement Learning with Deep ReLU Networks
 - Thanh Nguyen-Tang, Sunil Gupta, Hung Tran-The, and Svetha Venkatesh. arXiv, 2021.
 - Model Selection in Batch Policy Optimization
 - Jonathan N. Lee, George Tucker, Ofir Nachum, and Bo Dai. arXiv, 2021.
 - Learning Contraction Policies from Offline Data
 - Navid Rezazadeh, Maxwell Kolarich, Solmaz S. Kia, and Negar Mehr. arXiv, 2021.
 - CoMPS: Continual Meta Policy Search
 - Glen Berseth, Zhiwei Zhang, Grace Zhang, Chelsea Finn, Sergey Levine. arXiv, 2021.
 - MESA: Offline Meta-RL for Safe Adaptation and Fault Tolerance
 - Michael Luo, Ashwin Balakrishna, Brijen Thananjeyan, Suraj Nair, Julian Ibarz, Jie Tan, Chelsea Finn, Ion Stoica, and Ken Goldberg. arXiv, 2021.
 - Offline Pre-trained Multi-Agent Decision Transformer: One Big Sequence Model Conquers All StarCraftII Tasks
 - Linghui Meng, Muning Wen, Yaodong Yang, Chenyang Le, Xiyun Li, Weinan Zhang, Ying Wen, Haifeng Zhang, Jun Wang, and Bo Xu. arXiv, 2021.
 - Policy Gradient and Actor-Critic Learning in Continuous Time and Space: Theory and Algorithms
 - Yanwei Jia and Xun Yu Zhou. arXiv, 2021.
 - Offline Reinforcement Learning: Fundamental Barriers for Value Function Approximation [video]
-

-
- Dylan J. Foster, Akshay Krishnamurthy, David Simchi-Levi, and Yunzong Xu. arXiv, 2021.
 - UMBRELLA: Uncertainty-Aware Model-Based Offline Reinforcement Learning Leveraging Planning
 - Christopher Diehl, Timo Sievernich, Martin Krüger, Frank Hoffmann, and Torsten Bertran. arXiv, 2021.
 - Exploiting Action Impact Regularity and Partially Known Models for Offline Reinforcement Learning
 - Vincent Liu, James Wright, and Martha White. arXiv, 2021.
 - Batch Reinforcement Learning from Crowds
 - Guoxi Zhang and Hisashi Kashima. arXiv, 2021.
 - SCORE: Spurious CORrelation REDuction for Offline Reinforcement Learning
 - Zhihong Deng, Zuyue Fu, Lingxiao Wang, Zhuoran Yang, Chenjia Bai, Zhaoran Wang, and Jing Jiang. arXiv, 2021.
 - Safely Bridging Offline and Online Reinforcement Learning
 - Wanqiao Xu, Kan Xu, Hamsa Bastani, and Osbert Bastani. arXiv, 2021.
 - Efficient Robotic Manipulation Through Offline-to-Online Reinforcement Learning and Goal-Aware State Information
 - Jin Li, Xianyuan Zhan, Zixu Xiao, and Guyue Zhou. arXiv, 2021.
 - Value Penalized Q-Learning for Recommender Systems
 - Chengqian Gao, Ke Xu, and Peilin Zhao. arXiv, 2021.
 - Offline Reinforcement Learning with Soft Behavior Regularization
 - Haoran Xu, Xianyuan Zhan, Jianxiong Li, and Honglei Yin. arXiv, 2021.
 - Planning from Pixels in Environments with Combinatorially Hard Search Spaces
 - Marco Bagatella, Mirek Olšák, Michal Rolínek, and Georg Martius. arXiv, 2021.
 - StARformer: Transformer with State-Action-Reward Representations
 - Jinghuan Shang and Michael S. Ryoo. arXiv, 2021.
 - Offline RL With Resource Constrained Online Deployment [code]
 - Jayanth Reddy Regatti, Aniket Anand Deshmukh, Frank Cheng, Young Hun Jung, Abhishek Gupta, and Urun Dogan. arXiv, 2021.
-

-
- Lifelong Robotic Reinforcement Learning by Retaining Experiences [website]
 - Annie Xie and Chelsea Finn. arXiv, 2021.
 - Dual Behavior Regularized Reinforcement Learning
 - Chapman Siu, Jason Traish, and Richard Yi Da Xu. arXiv, 2021.
 - DCUR: Data Curriculum for Teaching via Samples with Reinforcement Learning [website] [code]
 - Daniel Seita, Abhinav Gopal, Zhao Mandi, and John Canny. arXiv, 2021.
 - DROMO: Distributionally Robust Offline Model-based Policy Optimization
 - Ruizhen Liu, Dazhi Zhong, and Zhicong Chen. arXiv, 2021.
 - Implicit Behavioral Cloning
 - Pete Florence, Corey Lynch, Andy Zeng, Oscar Ramirez, Ayzaan Wahid, Laura Downs, Adrian Wong, Johnny Lee, Igor Mordatch, and Jonathan Tompson. arXiv, 2021.
 - Reducing Conservativeness Oriented Offline Reinforcement Learning
 - Hongchang Zhang, Jianzhun Shao, Yuhang Jiang, Shuncheng He, and Xiangyang Ji. arXiv, 2021.
 - Policy Gradients Incorporating the Future
 - David Venuto, Elaine Lau, Doina Precup, and Ofir Nachum. arXiv, 2021.
 - Offline Decentralized Multi-Agent Reinforcement Learning
 - Jiechuan Jiang and Zongqing Lu. arXiv, 2021.
 - OPAL: Offline Preference-Based Apprenticeship Learning [website]
 - Daniel Shin and Daniel S. Brown. arXiv, 2021.
 - Constraints Penalized Q-Learning for Safe Offline Reinforcement Learning
 - Haoran Xu, Xianyuan Zhan, and Xiangyu Zhu. arXiv, 2021.
 - Where is the Grass Greener? Revisiting Generalized Policy Iteration for Offline Reinforcement Learning
 - Lionel Blondé and Alexandros Kalousis. arXiv, 2021.
 - The Least Restriction for Offline Reinforcement Learning
 - Zizhou Su. arXiv, 2021.
-

-
- Offline-to-Online Reinforcement Learning via Balanced Replay and Pessimistic Q-Ensemble
 - Seunghyun Lee, Younggyo Seo, Kimin Lee, Pieter Abbeel, and Jinwoo Shin. arXiv, 2021.
 - Causal Reinforcement Learning using Observational and Interventional Data
 - Maxime Gasse, Damien Grasset, Guillaume Gaudron, and Pierre-Yves Oudeyer. arXiv, 2021.
 - On the Sample Complexity of Batch Reinforcement Learning with Policy-Induced Data
 - Chenjun Xiao, Ilbin Lee, Bo Dai, Dale Schuurmans, and Csaba Szepesvari. arXiv, 2021.
 - Behavioral Priors and Dynamics Models: Improving Performance and Domain Transfer in Offline RL [website]
 - Catherine Cang, Aravind Rajeswaran, Pieter Abbeel, and Michael Laskin. arXiv, 2021.
 - On Multi-objective Policy Optimization as a Tool for Reinforcement Learning
 - Abbas Abdolmaleki, Sandy H. Huang, Giulia Vezzani, Bobak Shahriari, Jost Tobias Springenberg, Shruti Mishra, Dhruva TB, Arunkumar Byravan, Konstantinos Bousmalis, Andras Gyorgy, Csaba Szepesvari, Raia Hadsell, Nicolas Heess, and Martin Riedmiller. arXiv, 2021.
 - Offline Reinforcement Learning as Anti-Exploration
 - Shideh Rezaeifar, Robert Dadashi, Nino Vieillard, Léonard Hussenot, Olivier Bachem, Olivier Pietquin, and Matthieu Geist. arXiv, 2021.
 - Corruption-Robust Offline Reinforcement Learning
 - Xuezhou Zhang, Yiding Chen, Jerry Zhu, and Wen Sun. arXiv, 2021.
 - Offline Inverse Reinforcement Learning
 - Firas Jarboui and Vianney Perchet. arXiv, 2021.
 - Heuristic-Guided Reinforcement Learning
 - Ching-An Cheng, Andrey Kolobov, and Adith Swaminathan. arXiv, 2021.
 - Reinforcement Learning as One Big Sequence Modeling Problem
 - Michael Janner, Qiyang Li, and Sergey Levine. arXiv, 2021.
 - Decision Transformer: Reinforcement Learning via Sequence Modeling
 - Lili Chen, Kevin Lu, Aravind Rajeswaran, Kimin Lee, Aditya Grover, Michael Laskin, Pieter Abbeel, Aravind Srinivas, and Igor Mordatch. arXiv, 2021.
 - Model-Based Offline Planning with Trajectory Pruning

-
- Xianyuan Zhan, Xiangyu Zhu, and Haoran Xu. arXiv, 2021.
 - InferNet for Delayed Reinforcement Tasks: Addressing the Temporal Credit Assignment Problem
 - Markel Sanz Ausin, Hamoon Azizzoltani, Song Ju, Yeo Jin Kim, and Min Chi. arXiv, 2021.
 - Infinite-Horizon Offline Reinforcement Learning with Linear Function Approximation: Curse of Dimensionality and Algorithm [video]
 - Lin Chen, Bruno Scherrer, and Peter L. Bartlett. arXiv, 2021.
 - MT-Opt: Continuous Multi-Task Robotic Reinforcement Learning at Scale [website]
 - Dmitry Kalashnikov, Jacob Varley, Yevgen Chebotar, Benjamin Swanson, Rico Jonschkowski, Chelsea Finn, Sergey Levine, and Karol Hausman. arXiv, 2021.
 - Distributional Offline Continuous-Time Reinforcement Learning with Neural Physics-Informed PDEs (SciPhy RL for DOCTR-L)
 - Igor Halperin. arXiv, 2021.
 - Regularized Behavior Value Estimation
 - Caglar Gulcehre, Sergio Gómez Colmenarejo, Ziyu Wang, Jakub Sygnowski, Thomas Paine, Konrad Zolna, Yutian Chen, Matthew Hoffman, Razvan Pascanu, and Nando de Freitas. arXiv, 2021.
 - Improved Context-Based Offline Meta-RL with Attention and Contrastive Learning
 - Lanqing Li, Yuanhao Huang, and Dijun Luo. arXiv, 2021.
 - Instrumental Variable Value Iteration for Causal Offline Reinforcement Learning
 - Luofeng Liao, Zuyue Fu, Zhuoran Yang, Mladen Kolar, and Zhaoran Wang. arXiv, 2021.
 - GELATO: Geometrically Enriched Latent Model for Offline Reinforcement Learning
 - Guy Tennenholtz, Nir Baram, and Shie Mannor. arXiv, 2021.
 - MUSBO: Model-based Uncertainty Regularized and Sample Efficient Batch Optimization for Deployment Constrained Reinforcement Learning
 - DiJia Su, Jason D. Lee, John M. Mulvey, and H. Vincent Poor. arXiv, 2021.
 - Continuous Doubly Constrained Batch Reinforcement Learning
 - Rasool Fakoor, Jonas Mueller, Pratik Chaudhari, and Alexander J. Smola. arXiv, 2021.
 - Q-Value Weighted Regression: Reinforcement Learning with Limited Data
-

-
- Piotr Kozakowski, Łukasz Kaiser, Henryk Michalewski, Afroz Mohiuddin, and Katarzyna Kańska. arXiv, 2021.
 - Finite Sample Analysis of Minimax Offline Reinforcement Learning: Completeness, Fast Rates and First-Order Efficiency
 - Masatoshi Uehara, Masaaki Imaizumi, Nan Jiang, Nathan Kallus, Wen Sun, and Tengyang Xie. arXiv, 2021.
 - Fast Rates for the Regret of Offline Reinforcement Learning [video]
 - Yichun Hu, Nathan Kallus, and Masatoshi Uehara. arXiv, 2021.
 - Safe Policy Learning through Extrapolation: Application to Pre-trial Risk Assessment [video]
 - Eli Ben-Michael, D. James Greiner, Kosuke Imai, and Zhichao Jiang.
 - Weighted Model Estimation for Offline Model-based Reinforcement Learning
 - Toru Hishinuma and Kei Senda. NeurIPS, 2021.
 - A Minimalist Approach to Offline Reinforcement Learning
 - Scott Fujimoto and Shixiang Shane Gu. NeurIPS, 2021.
 - Conservative Offline Distributional Reinforcement Learning
 - Yecheng Jason Ma, Dinesh Jayaraman, and Osbert Bastani. NeurIPS, 2021.
 - Pessimism Meets Invariance: Provably Efficient Offline Mean-Field Multi-Agent RL
 - Minshuo Chen, Yan Li, Ethan Wang, Zhuoran Yang, Zhaoran Wang, and Tuo Zhao. NeurIPS, 2021.
 - Believe What You See: Implicit Constraint Approach for Offline Multi-Agent Reinforcement Learning
 - Yiqin Yang, Xiaoteng Ma, Chenghao Li, Zewu Zheng, Qiyuan Zhang, Gao Huang, Jun Yang, and Qianchuan Zhao. NeurIPS, 2021.
 - Provable Benefits of Actor-Critic Methods for Offline Reinforcement Learning
 - Andrea Zanette, Martin J. Wainwright, and Emma Brunskill. NeurIPS, 2021.
 - Multi-Objective SPIBB: Seldonian Offline Policy Improvement with Safety Constraints in Finite MDPs
 - Harsh Satija, Philip S. Thomas, Joelle Pineau, and Romain Laroche. NeurIPS, 2021.
 - Offline Reinforcement Learning as One Big Sequence Modeling Problem
-

-
- Michael Janner, Qiyang Li, and Sergey Levine. NeurIPS, 2021.
 - Bridging Offline Reinforcement Learning and Imitation Learning: A Tale of Pessimism [video]
 - Paria Rashidinejad, Banghua Zhu, Cong Ma, Jiantao Jiao, and Stuart Russell. NeurIPS, 2021.
 - Offline Reinforcement Learning with Reverse Model-based Imagination
 - Jianhao Wang, Wenzhe Li, Haozhe Jiang, Guangxiang Zhu, Siyuan Li, and Chongjie Zhang. NeurIPS, 2021.
 - Offline Meta Reinforcement Learning – Identifiability Challenges and Effective Data Collection Strategies
 - Ron Dorfman, Idan Shenfeld, and Aviv Tamar. NeurIPS, 2021.
 - Nearly Horizon-Free Offline Reinforcement Learning
 - Tongzheng Ren, Jialian Li, Bo Dai, Simon S. Du, and Sujay Sanghavi. NeurIPS, 2021.
 - Conservative Data Sharing for Multi-Task Offline Reinforcement Learning
 - Tianhe Yu, Aviral Kumar, Yevgen Chebotar, Karol Hausman, Sergey Levine, and Chelsea Finn. NeurIPS, 2021.
 - Online and Offline Reinforcement Learning by Planning with a Learned Model
 - Julian Schrittwieser, Thomas Hubert, Amol Mandhane, Mohammadamin Barekatain, Ioannis Antonoglou, and David Silver. NeurIPS, 2021.
 - Policy Finetuning: Bridging Sample-Efficient Offline and Online Reinforcement Learning
 - Tengyang Xie, Nan Jiang, Huan Wang, Caiming Xiong, and Yu Bai. NeurIPS, 2021.
 - Offline RL Without Off-Policy Evaluation
 - David Brandfonbrener, William F. Whitney, Rajesh Ranganath, and Joan Bruna. NeurIPS, 2021.
 - Offline Model-based Adaptable Policy Learning
 - Xiong-Hui Chen, Yang Yu, Qingyang Li, Fan-Ming Luo, Zhiwei Tony Qin, Shang Wenjie, and Jieping Ye. NeurIPS, 2021.
 - COMBO: Conservative Offline Model-Based Policy Optimization
 - Tianhe Yu, Aviral Kumar, Rafael Rafailov, Aravind Rajeswaran, Sergey Levine, and Chelsea Finn. NeurIPS, 2021.
-

-
- PerSim: Data-Efficient Offline Reinforcement Learning with Heterogeneous Agents via Personalized Simulators
 - Anish Agarwal, Abdullah Alomar, Varkey Alumootil, Devavrat Shah, Dennis Shen, Zhi Xu, and Cindy Yang. NeurIPS, 2021.
 - Near-Optimal Offline Reinforcement Learning via Double Variance Reduction
 - Ming Yin, Yu Bai, and Yu-Xiang Wang. NeurIPS, 2021.
 - Bellman-consistent Pessimism for Offline Reinforcement Learning [video]
 - Tengyang Xie, Ching-An Cheng, Nan Jiang, Paul Mineiro, and Alekh Agarwal. NeurIPS, 2021.
 - The Difficulty of Passive Learning in Deep Reinforcement Learning
 - Georg Ostrovski, Pablo Samuel Castro, and Will Dabney. NeurIPS, 2021.
 - Uncertainty-Based Offline Reinforcement Learning with Diversified Q-Ensemble
 - Gaon An, Seungyong Moon, Jang-Hyun Kim, and Hyun Oh Song. NeurIPS, 2021.
 - Towards Instance-Optimal Offline Reinforcement Learning with Pessimism
 - Ming Yin and Yu-Xiang Wang. NeurIPS, 2021.
 - EMaQ: Expected-Max Q-Learning Operator for Simple Yet Effective Offline and Online RL
 - Seyed Kamyar Seyed Ghasemipour, Dale Schuurmans, and Shixiang Shane Gu. ICML, 2021.
 - Actionable Models: Unsupervised Offline Reinforcement Learning of Robotic Skills [website]
 - Yevgen Chebotar, Karol Hausman, Yao Lu, Ted Xiao, Dmitry Kalashnikov, Jake Varley, Alex Irpan, Benjamin Eysenbach, Ryan Julian, Chelsea Finn, and Sergey Levine. ICML, 2021.
 - Is Pessimism Provably Efficient for Offline RL? [video]
 - Ying Jin, Zhuoran Yang, and Zhaoran Wang. ICML, 2021.
 - Representation Matters: Offline Pretraining for Sequential Decision Making
 - Mengjiao Yang and Ofir Nachum. ICML, 2021.
 - Offline Reinforcement Learning with Pseudometric Learning
 - Robert Dadashi, Shideh Rezaeifar, Nino Vieillard, Léonard Hussenot, Olivier Pietquin, and Matthieu Geist. ICML, 2021.
-

-
- Augmented World Models Facilitate Zero-Shot Dynamics Generalization From a Single Offline Environment
 - Philip J. Ball, Cong Lu, Jack Parker-Holder, and Stephen Roberts. ICML, 2021.
 - Offline Contextual Bandits with Overparameterized Models
 - David Brandfonbrener, William F. Whitney, Rajesh Ranganath and Joan Bruna. ICML, 2021.
 - Risk Bounds and Rademacher Complexity in Batch Reinforcement Learning
 - Yaqi Duan, Chi Jin, and Zhiyuan Li. ICML, 2021.
 - Offline Reinforcement Learning with Fisher Divergence Critic Regularization
 - Ilya Kostrikov, Jonathan Tompson, Rob Fergus, and Ofir Nachum. ICML, 2021.
 - OptiDICE: Offline Policy Optimization via Stationary Distribution Correction Estimation
 - Jongmin Lee, Wonseok Jeon, Byung-Jun Lee, Joelle Pineau, and Kee-Eung Kim. ICML, 2021.
 - Uncertainty Weighted Actor-Critic for Offline Reinforcement Learning
 - Yue Wu, Shuangfei Zhai, Nitish Srivastava, Joshua Susskind, Jian Zhang, Ruslan Salakhutdinov, and Hanlin Goh. ICML, 2021.
 - Vector Quantized Models for Planning
 - Sherjil Ozair, Yazhe Li, Ali Razavi, Ioannis Antonoglou, Aäron van den Oord, and Oriol Vinyals. ICML, 2021.
 - Exponential Lower Bounds for Batch Reinforcement Learning: Batch RL can be Exponentially Harder than Online RL [video]
 - Andrea Zanette. ICML, 2021.
 - Instabilities of Offline RL with Pre-Trained Neural Representation
 - Ruosong Wang, Yifan Wu, Ruslan Salakhutdinov, and Sham M. Kakade. ICML, 2021.
 - Offline Meta-Reinforcement Learning with Advantage Weighting
 - Eric Mitchell, Rafael Rafailov, Xue Bin Peng, Sergey Levine, and Chelsea Finn. ICML, 2021.
 - Model-Based Offline Planning [video]
 - Arthur Argenson and Gabriel Dulac-Arnold. ICLR, 2021.
 - Batch Reinforcement Learning Through Continuation Method

-
- Yijie Guo, Shengyu Feng, Nicolas Le Roux, Ed Chi, Honglak Lee, and Minmin Chen. ICLR, 2021.
 - Model-Based Visual Planning with Self-Supervised Functional Distances
 - Stephen Tian, Suraj Nair, Frederik Ebert, Sudeep Dasari, Benjamin Eysenbach, Chelsea Finn, and Sergey Levine. ICLR, 2021.
 - Deployment-Efficient Reinforcement Learning via Model-Based Offline Optimization
 - Tatsuya Matsushima, Hiroki Furuta, Yutaka Matsuo, Ofir Nachum, and Shixiang Gu. ICLR, 2021.
 - Efficient Fully-Offline Meta-Reinforcement Learning via Distance Metric Learning and Behavior Regularization
 - Lanqing Li, Rui Yang, and Dijun Luo. ICLR, 2021.
 - DeepAveragers: Offline Reinforcement Learning by Solving Derived Non-Parametric MDPs
 - Aayam Kumar Shrestha, Stefan Lee, Prasad Tadepalli, and Alan Fern. ICLR, 2021.
 - What are the Statistical Limits of Offline RL with Linear Function Approximation? [video]
 - Ruosong Wang, Dean Foster, and Sham M. Kakade. ICLR, 2021.
 - Reset-Free Lifelong Learning with Skill-Space Planning [website]
 - Kevin Lu, Aditya Grover, Pieter Abbeel, and Igor Mordatch. ICLR, 2021.
 - Risk-Averse Offline Reinforcement Learning
 - Núria Armengol Urpí, Sebastian Curi, and Andreas Krause. ICLR, 2021.
 - Finite-Sample Regret Bound for Distributionally Robust Offline Tabular Reinforcement Learning
 - Zhengqing Zhou, Zhengyuan Zhou, Qinxun Bai, Linhai Qiu, Jose Blanchet, and Peter Glynn. AISTATS, 2021.
 - Exploration by Maximizing Rényi Entropy for Reward-Free RL Framework
 - Chuheng Zhang, Yuanying Cai, Longbo Huang, and Jian Li. AAAI, 2021.
 - Efficient Self-Supervised Data Collection for Offline Robot Learning
 - Shadi Endrawis, Gal Leibovich, Guy Jacob, Gal Novik, Aviv Tamar. ICRA, 2021.
 - Boosting Offline Reinforcement Learning with Residual Generative Modeling
 - Hua Wei, Deheng Ye, Zhao Liu, Hao Wu, Bo Yuan, Qiang Fu, Wei Yang, and Zhenhui (Jessie)Li. IJCAI, 2021.
-

-
- BRAC+: Improved Behavior Regularized Actor Critic for Offline Reinforcement Learning
 - Chi Zhang, Sanmukh Rao Kuppannagari, and Viktor K Prasanna. ACML, 2021.
 - Behavior Constraining in Weight Space for Offline Reinforcement Learning
 - Phillip Swazinna, Steffen Udluft, Daniel Hein, and Thomas Runkler. ESANN, 2021.
 - Finite-Sample Analysis For Decentralized Batch Multi-Agent Reinforcement Learning With Networked Agents
 - Kaiqing Zhang, Zhuoran Yang, Han Liu, Tong Zhang, and Tamer Başar. IEEE T AUTOMATIC CONTROL, 2021.
 - Can Active Sampling Reduce Causal Confusion in Offline Reinforcement Learning?
 - Gunshi Gupta, Tim G. J. Rudner, Rowan Thomas McAllister, Adrien Gaidon, and Yarin Gal. CLear, 2021.
 - Reinforcement Learning via Fenchel-Rockafellar Duality [software]
 - Ofir Nachum and Bo Dai. arXiv, 2020.
 - AWAC: Accelerating Online Reinforcement Learning with Offline Datasets [website] [code] [blog]
 - Ashvin Nair, Abhishek Gupta, Murtaza Dalal, and Sergey Levine. arXiv, 2020.
 - Sparse Feature Selection Makes Batch Reinforcement Learning More Sample Efficient
 - Botao Hao, Yaqi Duan, Tor Lattimore, Csaba Szepesvári, and Mengdi Wang. arXiv, 2020.
 - A Variant of the Wang-Foster-Kakade Lower Bound for the Discounted Setting
 - Philip Amortila, Nan Jiang, and Tengyang Xie. arXiv, 2020.
 - Batch Reinforcement Learning with a Nonparametric Off-Policy Policy Gradient
 - Samuele Tosatto, João Carvalho, and Jan Peters. arXiv, 2020.
 - Batch Value-function Approximation with Only Realizability
 - Tengyang Xie and Nan Jiang. arXiv2020.
 - DRIFT: Deep Reinforcement Learning for Functional Software Testing
 - Luke Harries, Rebekah Storan Clarke, Timothy Chapman, Swamy V. P. L. N. Nallamalli, Levent Ozgur, Shuktika Jain, Alex Leung, Steve Lim, Aaron Dietrich, José Miguel Hernández-Lobato, Tom Ellis, Cheng Zhang, and Kamil Ciosek. arXiv, 2020.
 - Causality and Batch Reinforcement Learning: Complementary Approaches To Planning In Unknown Domains
-

-
- James Bannon, Brad Windsor, Wenbo Song, and Tao Li. arXiv, 2020.
 - Goal-conditioned Batch Reinforcement Learning for Rotation Invariant Locomotion [code]
 - Aditi Mavalankar. arXiv, 2020.
 - Semi-Supervised Reward Learning for Offline Reinforcement Learning
 - Ksenia Konyushkova, Konrad Zolna, Yusuf Aytar, Alexander Novikov, Scott Reed, Serkan Cabi, and Nando de Freitas. arXiv, 2020.
 - Sample-Efficient Reinforcement Learning via Counterfactual-Based Data Augmentation
 - Chaochao Lu, Biwei Huang, Ke Wang, José Miguel Hernández-Lobato, Kun Zhang, and Bernhard Schölkopf. arXiv, 2020.
 - Offline Reinforcement Learning from Images with Latent Space Models [website]
 - Rafael Rafailov, Tianhe Yu, Aravind Rajeswaran, and Chelsea Finn. arXiv, 2020.
 - POPO: Pessimistic Offline Policy Optimization
 - Qiang He and Xinwen Hou. arXiv, 2020.
 - Reinforcement Learning with Videos: Combining Offline Observations with Interaction
 - Karl Schmeckpeper, Oleh Rybkin, Kostas Daniilidis, Sergey Levine, and Chelsea Finn. arXiv, 2020.
 - Recovery RL: Safe Reinforcement Learning with Learned Recovery Zones [website]
 - Brijen Thananjeyan, Ashwin Balakrishna, Suraj Nair, Michael Luo, Krishnan Srinivasan, Minh Hwang, Joseph E. Gonzalez, Julian Ibarz, Chelsea Finn, and Ken Goldberg. arXiv, 2020.
 - Implicit Under-Parameterization Inhibits Data-Efficient Deep Reinforcement Learning
 - Aviral Kumar, Rishabh Agarwal, Dibya Ghosh, and Sergey Levine. arXiv, 2020.
 - OPAL: Offline Primitive Discovery for Accelerating Offline Reinforcement Learning [website]
 - Anurag Ajay, Aviral Kumar, Pulkit Agrawal, Sergey Levine, and Ofir Nachum. arXiv, 2020.
 - Batch Exploration with Examples for Scalable Robotic Reinforcement Learning
 - Annie S. Chen, HyunJi Nam, Suraj Nair, and Chelsea Finn. arXiv, 2020.
 - Learning Dexterous Manipulation from Suboptimal Experts [website]
 - Rae Jeong, Jost Tobias Springenberg, Jackie Kay, Daniel Zheng, Yuxiang Zhou, Alexandre Galashov, Nicolas Heess, and Francesco Nori. arXiv, 2020.
-

-
- The Reinforcement Learning-Based Multi-Agent Cooperative Approach for the Adaptive Speed Regulation on a Metallurgical Pickling Line
 - Anna Bogomolova, Kseniia Kingsep, and Boris Voskresenskii. arXiv, 2020.
 - Overcoming Model Bias for Robust Offline Deep Reinforcement Learning [dataset]
 - Phillip Swazinna, Steffen Udluft, and Thomas Runkler. arXiv, 2020.
 - Offline Meta Learning of Exploration
 - Ron Dorfman, Idan Shenfeld, and Aviv Tamar. arXiv, 2020.
 - EMaQ: Expected-Max Q-Learning Operator for Simple Yet Effective Offline and Online RL
 - Seyed Kamyar Seyed Ghasemipour, Dale Schuurmans, and Shixiang Shane Gu. arXiv, 2020.
 - Hyperparameter Selection for Offline Reinforcement Learning
 - Tom Le Paine, Cosmin Paduraru, Andrea Michi, Caglar Gulcehre, Konrad Zolna, Alexander Novikov, Ziyu Wang, and Nando de Freitas. arXiv, 2020.
 - Interpretable Control by Reinforcement Learning
 - Daniel Hein, Steffen Limmer, and Thomas A. Runkler. arXiv, 2020.
 - Efficient Evaluation of Natural Stochastic Policies in Offline Reinforcement Learning [code]
 - Nathan Kallus and Masatoshi Uehara. arXiv, 2020.
 - Accelerating Online Reinforcement Learning with Offline Datasets [website] [blog]
 - Ashvin Nair, Murtaza Dalal, Abhishek Gupta, and Sergey Levine. arXiv, 2020.
 - DisCor: Corrective Feedback in Reinforcement Learning via Distribution Correction [blog]
 - Aviral Kumar, Abhishek Gupta, and Sergey Levine. arXiv, 2020.
 - Critic Regularized Regression
 - Ziyu Wang, Alexander Novikov, Konrad Zolna, Josh S. Merel, Jost Tobias Springenberg, Scott E. Reed, Bobak Shahriari, Noah Siegel, Caglar Gulcehre, Nicolas Heess, and Nando de Freitas. NeurIPS, 2020
 - Provably Good Batch Off-Policy Reinforcement Learning Without Great Exploration
 - Yao Liu, Adith Swaminathan, Alekh Agarwal, and Emma Brunskill. NeurIPS, 2020.
 - Conservative Q-Learning for Offline Reinforcement Learning [website] [code] [blog]

-
- Aviral Kumar, Aurick Zhou, George Tucker, and Sergey Levine. NeurIPS, 2020.
 - BAIL: Best-Action Imitation Learning for Batch Deep Reinforcement Learning
 - Xinyue Chen, Zijian Zhou, Zheng Wang, Che Wang, Yanqiu Wu, and Keith Ross. NeurIPS, 2020.
 - MOPO: Model-based Offline Policy Optimization [code]
 - Tianhe Yu, Garrett Thomas, Lantao Yu, Stefano Ermon, James Y. Zou, Sergey Levine, Chelsea Finn, and Tengyu Ma. NeurIPS, 2020.
 - MOREL: Model-Based Offline Reinforcement Learning [podcast]
 - Rahul Kidambi, Aravind Rajeswaran, Praneeth Netrapalli, and Thorsten Joachims. NeurIPS, 2020.
 - Expert-Supervised Reinforcement Learning for Offline Policy Learning and Evaluation
 - Aaron Sonabend, Junwei Lu, Leo Anthony Celi, Tianxi Cai, and Peter Szolovits. NeurIPS, 2020.
 - Multi-task Batch Reinforcement Learning with Metric Learning
 - Jiachen Li, Quan Vuong, Shuang Liu, Minghua Liu, Kamil Ciosek, Henrik Christensen, and Hao Su. NeurIPS, 2020.
 - Counterfactual Data Augmentation using Locally Factored Dynamics [code]
 - Silviu Pitis, Elliot Creager, and Animesh Garg. NeurIPS, 2020.
 - On Reward-Free Reinforcement Learning with Linear Function Approximation
 - Ruosong Wang, Simon S. Du, Lin Yang, and Russ R. Salakhutdinov. NeurIPS, 2020.
 - Constrained Policy Improvement for Safe and Efficient Reinforcement Learning
 - Elad Sarafian, Aviv Tamar, and Sarit Kraus. IJCAI, 2020.
 - BRPO: Batch Residual Policy Optimization [code]
 - Sungryull Sohn, Yinlam Chow, Jayden Ooi, Ofir Nachum, Honglak Lee, Ed Chi, and Craig Boutilier. IJCAI, 2020.
 - Keep Doing What Worked: Behavior Modelling Priors for Offline Reinforcement Learning
 - Noah Siegel, Jost Tobias Springenberg, Felix Berkenkamp, Abbas Abdolmaleki, Michael Neunert, Thomas Lampe, Roland Hafner, Nicolas Heess, and Martin Riedmiller. ICLR, 2020.

-
- COG: Connecting New Skills to Past Experience with Offline Reinforcement Learning [website] [blog] [code]
 - Avi Singh, Albert Yu, Jonathan Yang, Jesse Zhang, Aviral Kumar, and Sergey Levine. CoRL, 2020.
 - Accelerating Reinforcement Learning with Learned Skill Priors
 - Karl Pertsch, Youngwoon Lee, and Joseph J. Lim. CoRL, 2020.
 - PLAS: Latent Action Space for Offline Reinforcement Learning [website] [code]
 - Wenxuan Zhou, Sujay Bajracharya, and David Held. CoRL, 2020.
 - Scaling data-driven robotics with reward sketching and batch reinforcement learning [website]
 - Serkan Cabi, Sergio Gómez Colmenarejo, Alexander Novikov, Ksenia Konyushkova, Scott Reed, Rae Jeong, Konrad Zolna, Yusuf Aytar, David Budden, Mel Vecerik, Oleg Sushkov, David Barker, Jonathan Scholz, Misha Denil, Nando de Freitas, and Ziyu Wang. RSS, 2020.
 - Quantile QT-Opt for Risk-Aware Vision-Based Robotic Grasping
 - Cristian Bodnar, Adrian Li, Karol Hausman, Peter Pastor, and Mrinal Kalakrishnan. RSS, 2020.
 - Batch-Constrained Reinforcement Learning for Dynamic Distribution Network Reconfiguration
 - Yuanqi Gao, Wei Wang, Jie Shi, and Nanpeng Yu. IEEE T SMART GRID, 2020.
 - Behavior Regularized Offline Reinforcement Learning
 - Yifan Wu, George Tucker, and Ofir Nachum. arXiv, 2019.
 - Off-Policy Policy Gradient Algorithms by Constraining the State Distribution Shift
 - Riashat Islam, Komal K. Teru, Deepak Sharma, and Joelle Pineau. arXiv, 2019.
 - Advantage-Weighted Regression: Simple and Scalable Off-Policy Reinforcement Learning
 - Xue Bin Peng, Aviral Kumar, Grace Zhang, and Sergey Levine. arXiv, 2019.
 - AlgaeDICE: Policy Gradient from Arbitrary Experience
 - Ofir Nachum, Bo Dai, Ilya Kostrikov, Yinlam Chow, Lihong Li, and Dale Schuurmans. arXiv, 2019.
 - Stabilizing Off-Policy Q-Learning via Bootstrapping Error Reduction [website] [blog] [code]
 - Aviral Kumar, Justin Fu, George Tucker, and Sergey Levine. NeurIPS, 2019.
-

-
- Off-Policy Deep Reinforcement Learning without Exploration
 - Scott Fujimoto, David Meger, and Doina Precup. ICML, 2019.
 - Safe Policy Improvement with Baseline Bootstrapping
 - Romain Laroche, Paul Trichelair, and Remi Tachet Des Combes. ICML, 2019.
 - Information-Theoretic Considerations in Batch Reinforcement Learning
 - Jinglin Chen and Nan Jiang. ICML, 2019.
 - Batch Recurrent Q-Learning for Backchannel Generation Towards Engaging Agents
 - Nusrah Hussain, Engin Erzin, T. Metin Sezgin, and Yucel Yemez. ACII, 2019.
 - Safe Policy Improvement with Soft Baseline Bootstrapping
 - Kimia Nadjahi, Romain Laroche, and Rémi Tachet des Combes. ECML, 2019.
 - Importance Weighted Transfer of Samples in Reinforcement Learning
 - Andrea Tirinzoni, Andrea Sessa, Matteo Pirodda, and Marcello Restelli. ICML, 2018.
 - Scalable Deep Reinforcement Learning for Vision-Based Robotic Manipulation [website]
 - Dmitry Kalashnikov, Alex Irpan, Peter Pastor, Julian Ibarz, Alexander Herzog, Eric Jang, Deirdre Quillen, Ethan Holly, Mrinal Kalakrishnan, Vincent Vanhoucke, and Sergey Levine. CoRL, 2018.
 - Off-Policy Policy Gradient with State Distribution Correction
 - Yao Liu, Adith Swaminathan, Alekh Agarwal, and Emma Brunskill. UAI, 2018.
 - Behavioral Cloning from Observation
 - Faraz Torabi, Garrett Warnell, and Peter Stone. IJCAI, 2018.
 - Diverse Exploration for Fast and Safe Policy Improvement
 - Andrew Cohen, Lei Yu, and Robert Wright. AAAI, 2018.
 - Deep Exploration via Bootstrapped DQN
 - Ian Osband, Charles Blundell, Alexander Pritzel, and Benjamin Van Roy. NeurIPS, 2016.
 - Safe Policy Improvement by Minimizing Robust Baseline Regret
 - Mohammad Ghavamzadeh, Marek Petrik, and Yinlam Chow. NeurIPS, 2016.
 - Residential Demand Response Applications Using Batch Reinforcement Learning

-
- Frederik Ruelens, Bert Claessens, Stijn Vandael, Bart De Schutter, Robert Babuska, and Ronnie Belmans. arXiv, 2015.
 - Structural Return Maximization for Reinforcement Learning
 - Joshua Joseph, Javier Velez, and Nicholas Roy. arXiv, 2014.
 - Simultaneous Perturbation Algorithms for Batch Off-Policy Search
 - Raphael Fonteneau, and L.A. Prashanth. CDC, 2014.
 - Guided Policy Search
 - Sergey Levine, and Vladlen Koltun. ICML, 2013.
 - Off-Policy Actor-Critic
 - Thomas Degris, Martha White, and Richard S. Sutton. ICML, 2012.
 - PAC-Bayesian Policy Evaluation for Reinforcement Learning
 - Mahdi Mllani Fard, Joelle Pineau, and Csaba Szepesvari. UAI, 2011.
 - Tree-Based Batch Mode Reinforcement Learning
 - Damien Ernst, Pierre Geurts, and Louis Wehenkel. JMLR, 2005.
 - Neural Fitted Q Iteration–First Experiences with a Data Efficient Neural Reinforcement Learning Method
 - Martin Riedmiller. ECML, 2005.
 - Off-Policy Temporal-Difference Learning with Function Approximation
 - Doina Precup, Richard S. Sutton, and Sanjoy Dasgupta. ICML, 2001.

Offline RL: Benchmarks/Experiments

- ORL-AUDITOR: Dataset Auditing in Offline Deep Reinforcement Learning
 - Linkang Du, Min Chen, Mingyang Sun, Shouling Ji, Peng Cheng, Jiming Chen, and Zhikun Zhang. NDSS, 2024.
- Pearl: A Production-ready Reinforcement Learning Agent
 - Zheqing Zhu, Rodrigo de Salvo Braz, Jalaj Bhandari, Daniel Jiang, Yi Wan, Yonathan Efroni, Liyuan Wang, Ruiyang Xu, Hongbo Guo, Alex Nikulkov, Dmytro Korenkevych, Urun Dogan, Frank Cheng, Zheng Wu, and Wanqiao Xu. arXiv, 2023.

-
- LMRL Gym: Benchmarks for Multi-Turn Reinforcement Learning with Language Models
 - Marwa Abdulhai, Isadora White, Charlie Snell, Charles Sun, Joey Hong, Yuexiang Zhai, Kelvin Xu, and Sergey Levine. arXiv, 2023.
 - Robotic Manipulation Datasets for Offline Compositional Reinforcement Learning
 - Marcel Hussing, Jorge A. Mendez, Anisha Singrodia, Cassandra Kent, and Eric Eaton. arXiv, 2023.
 - Datasets and Benchmarks for Offline Safe Reinforcement Learning
 - Zuxin Liu, Zijian Guo, Haohong Lin, Yihang Yao, Jiacheng Zhu, Zhepeng Cen, Hanjiang Hu, Wenhao Yu, Tingnan Zhang, Jie Tan, and Ding Zhao. arXiv, 2023.
 - Improving and Benchmarking Offline Reinforcement Learning Algorithms
 - Bingyi Kang, Xiao Ma, Yirui Wang, Yang Yue, and Shuicheng Yan. arXiv, 2023.
 - Benchmarks and Algorithms for Offline Preference-Based Reward Learning
 - Daniel Shin, Anca D. Dragan, and Daniel S. Brown. arXiv, 2023.
 - Hokoff: Real Game Dataset from Honor of Kings and its Offline Reinforcement Learning Benchmarks
 - Yun Qu, Boyuan Wang, Jianzhun Shao, Yuhang Jiang, Chen Chen, Zhenbin Ye, Liu Linc, Yang Feng, Lin Lai, Hongyang Qin, Minwen Deng, Juchao Zhuo, Deheng Ye, Qiang Fu, Yang Guang, Wei Yang, Lanxiao Huang, and Xiangyang Ji. NeurIPS, 2023.
 - CORL: Research-oriented Deep Offline Reinforcement Learning Library [code]
 - Denis Tarasov, Alexander Nikulin, Dmitry Akimov, Vladislav Kurenkov, and Sergey Kolesnikov. NeurIPS, 2023.
 - Benchmarking Offline Reinforcement Learning on Real-Robot Hardware [dataset]
 - Nico Gürtler, Sebastian Blaes, Pavel Kolev, Felix Widmaier, Manuel Wuthrich, Stefan Bauer, Bernhard Schölkopf, and Georg Martius. ICLR, 2023.
 - Train Offline, Test Online: A Real Robot Learning Benchmark
 - Gaoyue Zhou, Victoria Dean, Mohan Kumar Srirama, Aravind Rajeswaran, Jyothish Pari, Kyle Hatch, Aryan Jain, Tianhe Yu, Pieter Abbeel, Lerrel Pinto, Chelsea Finn, and Abhinav Gupta. ICRA, 2023.
 - Benchmarking Offline Reinforcement Learning Algorithms for E-Commerce Order Fraud Evaluation

-
- Soysal Degirmenci and Chris Jones. arXiv, 2022.
 - Real World Offline Reinforcement Learning with Realistic Data Source [website] [dataset]
 - Gaoyue Zhou, Liyiming Ke, Siddhartha Srinivasa, Abhinav Gupta, Aravind Rajeswaran, and Vikash Kumar. arXiv, 2022.
 - Mind Your Data! Hiding Backdoors in Offline Reinforcement Learning Datasets
 - Chen Gong, Zhou Yang, Yunpeng Bai, Junda He, Jieke Shi, Arunesh Sinha, Bowen Xu, Xinwen Hou, Guoliang Fan, and David Lo. arXiv, 2022.
 - B2RL: An open-source Dataset for Building Batch Reinforcement Learning
 - Hsin-Yu Liu, Xiaohan Fu, Bharathan Balaji, Rajesh Gupta, and Dezhi Hong. arXiv, 2022.
 - An Empirical Study of Implicit Regularization in Deep Offline RL
 - Caglar Gulcehre, Srivatsan Srinivasan, Jakub Sygnowski, Georg Ostrovski, Mehrdad Farajtabar, Matt Hoffman, Razvan Pascanu, and Arnaud Doucet. arXiv, 2022.
 - Challenges and Opportunities in Offline Reinforcement Learning from Visual Observations
 - Cong Lu, Philip J. Ball, Tim G. J. Rudner, Jack Parker-Holder, Michael A. Osborne, and Yee Whye Teh. arXiv, 2022.
 - Don't Change the Algorithm, Change the Data: Exploratory Data for Offline Reinforcement Learning [code]
 - Denis Yarats, David Brandfonbrener, Hao Liu, Michael Laskin, Pieter Abbeel, Alessandro Lazaric, and Lerrel Pinto. arXiv, 2022.
 - The Challenges of Exploration for Offline Reinforcement Learning
 - Nathan Lambert, Markus Wulfmeier, William Whitney, Arunkumar Byravan, Michael Bloesch, Vibhavari Dasagi, Tim Hertweck, and Martin Riedmiller. arXiv, 2022.
 - Offline Equilibrium Finding [code]
 - Shuxin Li, Xinrun Wang, Jakub Cerny, Youzhi Zhang, Hau Chan, and Bo An. arXiv, 2022.
 - Comparing Model-free and Model-based Algorithms for Offline Reinforcement Learning
 - Phillip Swazinna, Steffen Udluft, Daniel Hein, and Thomas Runkler. arXiv, 2022.
 - Data-Efficient Pipeline for Offline Reinforcement Learning with Limited Data
 - Allen Nie, Yannis Flet-Berliac, Deon R. Jordan, William Steenbergen, and Emma Brunskill. NeurIPS, 2022.
-

-
- Dungeons and Data: A Large-Scale NetHack Dataset
 - Eric Hambro, Roberta Raileanu, Danielle Rothermel, Vegard Mella, Tim Rocktäschel, Heinrich Küttler, and Naila Murray. NeurIPS, 2022.
 - NeoRL: A Near Real-World Benchmark for Offline Reinforcement Learning [website] [code]
 - Rongjun Qin, Songyi Gao, Xingyuan Zhang, Zhen Xu, Shengkai Huang, Zewen Li, Weinan Zhang, and Yang Yu. NeurIPS, 2022.
 - A Closer Look at Offline RL Agents
 - Yuwei Fu, Di Wu and Benoit Boulet. NeurIPS, 2022.
 - Beyond Rewards: a Hierarchical Perspective on Offline Multiagent Behavioral Analysis
 - Shayegan Omidshafiei, Andrei Kapishnikov, Yannick Assogba, Lucas Dixon, and Been Kim. NeurIPS, 2022.
 - On the Effect of Pre-training for Transformer in Different Modality on Offline Reinforcement Learning
 - Shiro Takagi. NeurIPS, 2022.
 - Showing Your Offline Reinforcement Learning Work: Online Evaluation Budget Matters
 - Vladislav Kurenkov and Sergey Kolesnikov. ICML, 2022.
 - d3rlpy: An Offline Deep Reinforcement Learning Library [software]
 - Takuma Seno and Michita Imai. JMLR, 2022.
 - Understanding the Effects of Dataset Characteristics on Offline Reinforcement Learning [code]
 - Kajetan Schweighofer, Markus Hofmarcher, Marius-Constantin Dinu, Philipp Renz, Angela Bitto-Nemling, Vihang Patil, and Sepp Hochreiter. arXiv, 2021.
 - Interpretable performance analysis towards offline reinforcement learning: A dataset perspective
 - Chenyang Xi, Bo Tang, Jiajun Shen, Xinfu Liu, Feiyu Xiong, and Xueying Li. arXiv, 2021.
 - Comparison and Unification of Three Regularization Methods in Batch Reinforcement Learning
 - Sarah Rathnam, Susan A. Murphy, and Finale Doshi-Velez. arXiv, 2021.
 - RLDS: an Ecosystem to Generate, Share and Use Datasets in Reinforcement Learning [code]

-
- Sabela Ramos, Sertan Girgin, Léonard Hussenot, Damien Vincent, Hanna Yakubovich, Daniel Toyama, Anita Gergely, Piotr Stanczyk, Raphael Marinier, Jeremiah Harmsen, Olivier Pietquin, and Nikola Momchev. NeurIPS, 2021.
 - Measuring Data Quality for Dataset Selection in Offline Reinforcement Learning
 - Phillip Swazinna, Steffen Udluft, and Thomas Runkler. IEEE SSCI, 2021.
 - Offline Reinforcement Learning Hands-On
 - Louis Monier, Jakub Kmec, Alexandre Laterre, Thomas Pierrot, Valentin Courgeau, Olivier Sigaud, Karim Beguir. arXiv, 2020.
 - D4RL: Datasets for Deep Data-Driven Reinforcement Learning [website] [blog] [code]
 - Justin Fu, Aviral Kumar, Ofir Nachum, George Tucker, and Sergey Levine. arXiv, 2020.
 - RL Unplugged: Benchmarks for Offline Reinforcement Learning [code] [dataset]
 - Caglar Gulcehre, Ziyu Wang, Alexander Novikov, Tom Le Paine, Sergio Gomez Colmenarejo, Konrad Zolna, Rishabh Agarwal, Josh Merel, Daniel Mankowitz, Cosmin Paduraru, Gabriel Dulac-Arnold, Jerry Li, Mohammad Norouzi, Matt Hoffman, Ofir Nachum, George Tucker, Nicolas Heess, and Nando de Freitas. NeurIPS, 2020.
 - Benchmarking Batch Deep Reinforcement Learning Algorithms
 - Scott Fujimoto, Edoardo Conti, Mohammad Ghavamzadeh, and Joelle Pineau. arXiv, 2019.

Offline RL: Applications

- MOTO: Offline Pre-training to Online Fine-tuning for Model-based Robot Learning
 - Rafael Rafailov, Kyle Hatch, Victor Kolev, John D. Martin, Mariano Phielipp, and Chelsea Finn. arXiv, 2024.
- P2DT: Mitigating Forgetting in task-incremental Learning with progressive prompt Decision Transformer
 - Zhiyuan Wang, Xiaoyang Qu, Jing Xiao, Bokui Chen, and Jianzong Wang. ICASSP, 2024.
- Online Symbolic Music Alignment with Offline Reinforcement Learning
 - Silvan David Peter. arXiv, 2023.
- Advancing RAN Slicing with Offline Reinforcement Learning
 - Kun Yang, Shu-ping Yeh, Menglei Zhang, Jerry Sydir, Jing Yang, and Cong Shen. arXiv, 2023.

-
- Traffic Signal Control Using Lightweight Transformers: An Offline-to-Online RL Approach
 - Xingshuai Huang, Di Wu, and Benoit Boulet. arXiv, 2023.
 - Self-Driving Telescopes: Autonomous Scheduling of Astronomical Observation Campaigns with Offline Reinforcement Learning
 - Franco Terranova, M. Voetberg, Brian Nord, and Amanda Pagul. arXiv, 2023.
 - A Fully Data-Driven Approach for Realistic Traffic Signal Control Using Offline Reinforcement Learning
 - Jianxiong Li, Shichao Lin, Tianyu Shi, Chujie Tian, Yu Mei, Jian Song, Xianyuan Zhan, and Ruimin Li. arXiv, 2023.
 - Offline Reinforcement Learning for Wireless Network Optimization with Mixture Datasets
 - Kun Yang, Cong Shen, Jing Yang, Shu-ping Yeh, and Jerry Sydir. arXiv, 2023.
 - STEER: Unified Style Transfer with Expert Reinforcement
 - Skyler Hallinan, Faeze Brahman, Ximing Lu, Jaehun Jung, Sean Welleck, and Yejin Choi. arXiv, 2023.
 - Zero-Shot Goal-Directed Dialogue via RL on Imagined Conversations
 - Joey Hong, Sergey Levine, and Anca Dragan. arXiv, 2023.
 - Robot Fine-Tuning Made Easy: Pre-Training Rewards and Policies for Autonomous Real-World Reinforcement Learning
 - Jingyun Yang, Max Sobol Mark, Brandon Vu, Archit Sharma, Jeannette Bohg, and Chelsea Finn. arXiv, 2023.
 - Offline Reinforcement Learning for Optimizing Production Bidding Policies
 - Dmytro Korenkevych, Frank Cheng, Artsiom Balakir, Alex Nikulkov, Lingnan Gao, Zhihao Cen, Zuobing Xu, and Zheqing Zhu. arXiv, 2023.
 - End-to-end Offline Reinforcement Learning for Glycemia Control
 - Tristan Beolet, Alice Adenis, Erik Hunecker, and Maxime Louis. arXiv, 2023.
 - Leveraging Optimal Transport for Enhanced Offline Reinforcement Learning in Surgical Robotic Environments
 - Maryam Zare, Parham M. Kebria, and Abbas Khosravi. arXiv, 2023.
 - Learning RL-Policies for Joint Beamforming Without Exploration: A Batch Constrained Off-Policy Approach
-

-
- Heasung Kim and Sravan Ankireddy. arXiv, 2023.
 - Uncertainty-Aware Decision Transformer for Stochastic Driving Environments
 - Zenan Li, Fan Nie, Qiao Sun, Fang Da, and Hang Zhao. arXiv, 2023.
 - Boosting Offline Reinforcement Learning for Autonomous Driving with Hierarchical Latent Skills
 - Zenan Li, Fan Nie, Qiao Sun, Fang Da, and Hang Zhao. arXiv, 2023.
 - Robotic Offline RL from Internet Videos via Value-Function Pre-Training
 - Chethan Bhateja, Derek Guo, Dibya Ghosh, Anikait Singh, Manan Tomar, Quan Vuong, Yevgen Chebotar, Sergey Levine, and Aviral Kumar. arXiv, 2023.
 - VAPOR: Holonomic Legged Robot Navigation in Outdoor Vegetation Using Offline Reinforcement Learning
 - Kasun Weerakoon, Adarsh Jagan Sathyamoorthy, Mohamed Elnoor, and Dinesh Manocha. arXiv, 2023.
 - RLSynC: Offline-Online Reinforcement Learning for Synthon Completion
 - Frazier N. Baker, Ziqi Chen, and Xia Ning. arXiv, 2023.
 - Real Robot Challenge 2022: Learning Dexterous Manipulation from Offline Data in the Real World
 - Nico Gürtler, Felix Widmaier, Cansu Sancaktar, Sebastian Blaes, Pavel Kolev, Stefan Bauer, Manuel Wüthrich, Markus Wulfmeier, Martin Riedmiller, Arthur Allshire, Qiang Wang, Robert McCarthy, Hangyeol Kim, Jongchan Baek Pohang, Woogyong Kwon, Shanliang Qian, Yasunori Toshimitsu, Mike Yan Michelis, Amirhossein Kazemipour, Arman Raayatsanati, Hehui Zheng, Barnabasa Gavin Cangan, Bernhard Schölkopf, and Georg Martius. arXiv, 2023.
 - Reinforced Self-Training (ReST) for Language Modeling
 - Caglar Gulcehre, Tom Le Paine, Srivatsan Srinivasan, Ksenia Konyushkova, Lotte Weerts, Abhishek Sharma, Aditya Siddhant, Alex Ahern, Miaosen Wang, Chenjie Gu, Wolfgang Macherey, Arnaud Doucet, Orhan Firat, and Nando de Freitas. arXiv, 2023.
 - Aligning Language Models with Offline Reinforcement Learning from Human Feedback
 - Jian Hu, Li Tao, June Yang, and Chandler Zhou. arXiv, 2023.
 - Integrating Offline Reinforcement Learning with Transformers for Sequential Recommendation
 - Xumei Xi, Yuke Zhao, Quan Liu, Liwen Ouyang, and Yang Wu. arXiv, 2023.
-

-
- Offline Skill Graph (OSG): A Framework for Learning and Planning using Offline Reinforcement Learning Skills
 - Ben-ya Halevy, Yehudit Aperstein, and Dotan Di Castro. arXiv, 2023.
 - Improving Offline RL by Blending Heuristics
 - Sinong Geng, Aldo Pacchiano, Andrey Kolobov, and Ching-An Cheng. arXiv, 2023.
 - IQL-TD-MPC: Implicit Q-Learning for Hierarchical Model Predictive Control
 - Rohan Chitnis, Yingchen Xu, Bobak Hashemi, Lucas Lehnert, Urun Dogan, Zheqing Zhu, and Olivier Delalleau. arXiv, 2023.
 - Robust Reinforcement Learning Objectives for Sequential Recommender Systems
 - Melissa Mozifian, Tristan Sylvain, Dave Evans, and Lili Meng. arXiv, 2023.
 - The Benefits of Being Distributional: Small-Loss Bounds for Reinforcement Learning
 - Kaiwen Wang, Kevin Zhou, Runzhe Wu, Nathan Kallus, and Wen Sun. arXiv, 2023.
 - PROTO: Iterative Policy Regularized Offline-to-Online Reinforcement Learning
 - Jianxiong Li, Xiao Hu, Haoran Xu, Jingjing Liu, Xianyuan Zhan, and Ya-Qin Zhang. arXiv, 2023.
 - Matrix Estimation for Offline Reinforcement Learning with Low-Rank Structure
 - Xumei Xi, Christina Lee Yu, and Yudong Chen. arXiv, 2023.
 - Offline Experience Replay for Continual Offline Reinforcement Learning
 - Sibor Gai, Donglin Wang, and Li He. arXiv, 2023.
 - Causal Decision Transformer for Recommender Systems via Offline Reinforcement Learning
 - Siyu Wang, Xiaocong Chen, Dietmar Jannach, and Lina Yao. arXiv, 2023.
 - Data Might be Enough: Bridge Real-World Traffic Signal Control Using Offline Reinforcement Learning
 - Liang Zhang and Jianming Deng. arXiv, 2023.
 - User Retention-oriented Recommendation with Decision Transformer
 - Kesen Zhao, Lixin Zou, Xiangyu Zhao, Maolin Wang, and Dawei Yin. arXiv, 2023.
 - Learning to Control Autonomous Fleets from Observation via Offline Reinforcement Learning

-
- Carolin Schmidt, Daniele Gammelli, Francisco Camara Pereira, and Filipe Rodrigues. arXiv, 2023.
 - INVICTUS: Optimizing Boolean Logic Circuit Synthesis via Synergistic Learning and Search
 - Animesh Basak Chowdhury, Marco Romanelli, Benjamin Tan, Ramesh Karri, and Siddharth Garg. arXiv, 2023.
 - Learning Vision-based Robotic Manipulation Tasks Sequentially in Offline Reinforcement Learning Settings
 - Sudhir Pratap Yadav, Rajendra Nagar, and Suril V. Shah. arXiv, 2023.
 - Winning Solution of Real Robot Challenge III
 - Qiang Wang, Robert McCarthy, David Cordova Bulens, and Stephen J. Redmond. arXiv, 2023.
 - Learning-based MPC from Big Data Using Reinforcement Learning
 - Shambhuraj Sawant, Akhil S Anand, Dirk Reinhardt, and Sebastien Gros. arXiv, 2023.
 - Offline Reinforcement Learning for Mixture-of-Expert Dialogue Management
 - Dhawal Gupta, Yinlam Chow, Aza Tulepbergenov, Mohammad Ghavamzadeh, and Craig Boutilier. NeurIPS, 2023.
 - Beyond Reward: Offline Preference-guided Policy Optimization
 - Yachen Kang, Diyu Shi, Jinxin Liu, Li He, and Donglin Wang. ICML, 2023.
 - DevFormer: A Symmetric Transformer for Context-Aware Device Placement
 - Haeyeon Kim, Minsu Kim, Federico Berto, Joungho Kim, and Jinkyoo Park. ICML, 2023.
 - On the Effectiveness of Offline RL for Dialogue Response Generation
 - Paloma Sodhi, Felix Wu, Ethan R. Elenberg, Kilian Q. Weinberger, and Ryan McDonald. ICML, 2023.
 - Bidirectional Learning for Offline Model-based Biological Sequence Design
 - Can Chen, Yingxue Zhang, Xue Liu, and Mark Coates. ICML, 2023.
 - ChiPFormer: Transferable Chip Placement via Offline Decision Transformer
 - Yao Lai, Jinxin Liu, Zhentao Tang, Bin Wang, Jianye Hao, and Ping Luo. ICML, 2023.
 - Semi-Offline Reinforcement Learning for Optimized Text Generation

-
- Changyu Chen, Xiting Wang, Yiqiao Jin, Victor Ye Dong, Li Dong, Jie Cao, Yi Liu, and Rui Yan. ICML, 2023.
 - Neural Constraint Satisfaction: Hierarchical Abstraction for Combinatorial Generalization in Object Rearrangement
 - Michael Chang, Alyssa L. Dayan, Franziska Meier, Thomas L. Griffiths, Sergey Levine, and Amy Zhang. ICLR, 2023.
 - Offline RL for Natural Language Generation with Implicit Language Q Learning
 - Charlie Snell, Ilya Kostrikov, Yi Su, Mengjiao Yang, and Sergey Levine. ICLR, 2023.
 - Action-Quantized Offline Reinforcement Learning for Robotic Skill Learning
 - Jianlan Luo, Perry Dong, Jeffrey Wu, Aviral Kumar, Xinyang Geng, and Sergey Levine. CoRL, 2023.
 - Building Persona Consistent Dialogue Agents with Offline Reinforcement Learning
 - Ryan Shea and Zhou Yu. EMNLP, 2023.
 - Dialog Action-Aware Transformer for Dialog Policy Learning
 - Huimin Wang, Wai-Chung Kwan, and Kam-Fai Wong. SIGdial, 2023.
 - Can Offline Reinforcement Learning Help Natural Language Understanding?
 - Ziqi Zhang, Yile Wang, Yue Zhang, and Donglin Wang. arXiv, 2022.
 - NeurIPS 2022 Competition: Driving SMARTS
 - Amir Rasouli, Randy Goebel, Matthew E. Taylor, Iuliia Kotseruba, Soheil Alizadeh, Tianpei Yang, Montgomery Alban, Florian Shkurti, Yuzheng Zhuang, Adam Scibior, Kasra Rezaee, Animesh Garg, David Meger, Jun Luo, Liam Paull, Weinan Zhang, Xinyu Wang, and Xi Chen. arXiv, 2022.
 - Controlling Commercial Cooling Systems Using Reinforcement Learning
 - Jerry Luo, Cosmin Paduraru, Octavian Voicu, Yuri Chervonyi, Scott Munns, Jerry Li, Crystal Qian, Praneet Dutta, Jared Quincy Davis, Ningjia Wu, Xingwei Yang, Chu-Ming Chang, Ted Li, Rob Rose, Mingyan Fan, Hootan Nakhost, Tinglin Liu, Brian Kirkman, Frank Altamura, Lee Cline, Patrick Tonker, Joel Gouker, Dave Uden, Warren Buddy Bryan, Jason Law, Deeni Fatiha, Neil Satra, Juliet Rothenberg, Molly Carlin, Satish Tallapaka, Sims Witherspoon, David Parish, Peter Dolan, Chenyu Zhao, and Daniel J. Mankowitz.
 - Pre-Training for Robots: Offline RL Enables Learning New Tasks from a Handful of Trials [code]

-
- Aviral Kumar, Anikait Singh, Frederik Ebert, Yanlai Yang, Chelsea Finn, and Sergey Levine. arXiv, 2022.
 - Towards Safe Mechanical Ventilation Treatment Using Deep Offline Reinforcement Learning
 - Flemming Kondrup, Thomas Jiralerspong, Elaine Lau, Nathan de Lara, Jacob Shkrob, My Duc Tran, Doina Precup, and Sumana Basu. IAAI, 2023.
 - Learning-to-defer for sequential medical decision-making under uncertainty
 - Shalmali Joshi, Sonali Parbhoo, and Finale Doshi-Velez. TMLR, 2023.
 - Imitation Is Not Enough: Robustifying Imitation with Reinforcement Learning for Challenging Driving Scenarios
 - Yiren Lu, Justin Fu, George Tucker, Xinlei Pan, Eli Bronstein, Rebecca Roelofs, Benjamin Sapp, Brandyn White, Aleksandra Faust, Shimon Whiteson, Dragomir Anguelov, and Sergey Levine. arXiv, 2022.
 - Dialogue Evaluation with Offline Reinforcement Learning
 - Nurul Lubis, Christian Geishauser, Hsien-Chin Lin, Carel van Niekerk, Michael Heck, Shutong Feng, and Milica Gašić. arXiv, 2022.
 - Multi-Task Fusion via Reinforcement Learning for Long-Term User Satisfaction in Recommender Systems
 - Qihua Zhang, Junning Liu, Yuzhuo Dai, Yiyang Qi, Yifan Yuan, Kunlun Zheng, Fan Huang, and Xianfeng Tan. arXiv, 2022.
 - A Maintenance Planning Framework using Online and Offline Deep Reinforcement Learning
 - Zaharah A. Bukhsh, Nils Jansen, and Hajo Molegraaf. arXiv, 2022.
 - BCRLSP: An Offline Reinforcement Learning Framework for Sequential Targeted Promotion
 - Fanglin Chen, Xiao Liu, Bo Tang, Feiyu Xiong, Serim Hwang, and Guomian Zhuang. arXiv, 2022.
 - Learning Optimal Treatment Strategies for Sepsis Using Offline Reinforcement Learning in Continuous Space
 - Zeyu Wang, Huiying Zhao, Peng Ren, Yuxi Zhou, and Ming Sheng. arXiv, 2022.
 - Rethinking Reinforcement Learning for Recommendation: A Prompt Perspective
 - Xin Xin, Tiago Pimentel, Alexandros Karatzoglou, Pengjie Ren, Konstantina Christakopoulou, and Zhaochun Ren. arXiv, 2022.
-

-
- ARLO: A Framework for Automated Reinforcement Learning
 - Marco Mussi, Davide Lombarda, Alberto Maria Metelli, Francesco Trovò, and Marcello Restelli. arXiv, 2022.
 - A Reinforcement Learning-based Volt-VAR Control Dataset and Testing Environment
 - Yuanqi Gao and Nanpeng Yu. arXiv, 2022.
 - CHAI: A CHatbot AI for Task-Oriented Dialogue with Offline Reinforcement Learning
 - Siddharth Verma, Justin Fu, Mengjiao Yang, and Sergey Levine. arXiv, 2022.
 - Offline Reinforcement Learning for Safer Blood Glucose Control in People with Type 1 Diabetes [code]
 - Harry Emerson, Matt Guy, and Ryan McConville. arXiv, 2022.
 - CIRS: Bursting Filter Bubbles by Counterfactual Interactive Recommender System [code]
 - Chongming Gao, Wenqiang Lei, Jiawei Chen, Shiqi Wang, Xiangnan He, Shijun Li, Biao Li, Yuan Zhang, and Peng Jiang. arXiv, 2022.
 - A Conservative Q-Learning approach for handling distribution shift in sepsis treatment strategies
 - Pramod Kaushik, Sneha Kummetha, Perusha Moodley, and Raju S. Bapi. arXiv, 2022.
 - Optimizing Trajectories for Highway Driving with Offline Reinforcement Learning
 - Branka Mirchevska, Moritz Werling, and Joschka Boedecker. arXiv, 2022.
 - Offline Deep Reinforcement Learning for Dynamic Pricing of Consumer Credit
 - Raad Khraishi and Ramin Okhrati. arXiv, 2022.
 - Offline Reinforcement Learning for Mobile Notifications
 - Yiping Yuan, Ajith Muralidharan, Preetam Nandy, Miao Cheng, and Prkruthi Prabhakar. arXiv, 2022.
 - Offline Reinforcement Learning for Road Traffic Control
 - Mayuresh Kunjir and Sanjay Chawla. arXiv, 2022.
 - Sustainable Online Reinforcement Learning for Auto-bidding
 - Zhiyu Mou, Yusen Huo, Rongquan Bai, Mingzhou Xie, Chuan Yu, Jian Xu, and Bo Zheng. NeurIPS, 2022.

-
- Leveraging Factored Action Spaces for Efficient Offline Reinforcement Learning in Healthcare
 - Shengpu Tang, Maggie Makar, Michael W. Sjoding, Finale Doshi-Velez, and Jenna Wiens. NeurIPS, 2022.
 - Multi-objective Optimization of Notifications Using Offline Reinforcement Learning
 - Prakruthi Prabhakar, Yiping Yuan, Guangyu Yang, Wensheng Sun, and Ajith Muralidharan. KDD, 2022.
 - Pessimism meets VCG: Learning Dynamic Mechanism Design via Offline Reinforcement Learning
 - Boxiang Lyu, Zhaoran Wang, Mladen Kolar, and Zhuoran Yang. ICML, 2022.
 - GPT-Critic: Offline Reinforcement Learning for End-to-End Task-Oriented Dialogue Systems
 - Youngsoo Jang, Jongmin Lee, and Kee-Eung Kim. ICLR, 2022.
 - Offline Reinforcement Learning for Visual Navigation
 - Dhruv Shah, Arjun Bhorkar, Hrish Leen, Ilya Kostrikov, Nick Rhinehart, and Sergey Levine. CoRL, 2022.
 - Semi-Markov Offline Reinforcement Learning for Healthcare
 - Mehdi Fatemi, Mary Wu, Jeremy Petch, Walter Nelson, Stuart J. Connolly, Alexander Benz, Anthony Carnicelli, and Marzyeh Ghassemi. CHIL, 2022.
 - Automate Page Layout Optimization: An Offline Deep Q-Learning Approach
 - Zhou Qin and Wenyang Liu. RecSys, 2022.
 - RL4RS: A Real-World Benchmark for Reinforcement Learning based Recommender System [code] [dataset]
 - Kai Wang, Zhene Zou, Yue Shang, Qilin Deng, Minghao Zhao, Yile Liang, Runze Wu, Jianrong Tao, Xudong Shen, Tangjie Lyu, and Changjie Fan. arXiv, 2021.
 - Compressive Features in Offline Reinforcement Learning for Recommender Systems
 - Hung Nguyen, Minh Nguyen, Long Pham, and Jennifer Adorno Nieves. arXiv, 2021.
 - Causal-aware Safe Policy Improvement for Task-oriented dialogue
 - Govardana Sachithanandam Ramachandran, Kazuma Hashimoto, and Caiming Xiong. arXiv, 2021.
 - Offline Contextual Bandits for Wireless Network Optimization

-
- Miguel Suau, Alexandros Agapitos, David Lynch, Derek Farrell, Mingqi Zhou, and Aleksandar Milenovic. arXiv, 2021.
 - Identifying Decision Points for Safe and Interpretable Reinforcement Learning in Hypotension Treatment
 - Kristine Zhang, Yuanheng Wang, Jianzhun Du, Brian Chu, Leo Anthony Celi, Ryan Kindle, and Finale Doshi-Velez. arXiv, 2021.
 - Offline Reinforcement Learning for Autonomous Driving with Safety and Exploration Enhancement
 - Tianyu Shi, Dong Chen, Kaian Chen, and Zhaojian Li. arXiv, 2021.
 - Medical Dead-ends and Learning to Identify High-risk States and Treatments
 - Mehdi Fatemi, Taylor W. Killian, Jayakumar Subramanian, and Marzyeh Ghassemi. arXiv, 2021.
 - An Offline Deep Reinforcement Learning for Maintenance Decision-Making
 - Hamed Khorasgani, Haiyan Wang, Chetan Gupta, and Ahmed Farahat. arXiv, 2021.
 - Learning Language-Conditioned Robot Behavior from Offline Data and Crowd-Sourced Annotation
 - Suraj Nair, Eric Mitchell, Kevin Chen, Brian Ichter, Silvio Savarese, and Chelsea Finn. arXiv, 2021.
 - Offline-Online Reinforcement Learning for Energy Pricing in Office Demand Response: Lowering Energy and Data Costs
 - Doseok Jang, Lucas Spangher, Manan Khattar, Utkarsha Agwan, Selvaprabuh Nadarajah, and Costas Spanos. arXiv, 2021.
 - Offline reinforcement learning with uncertainty for treatment strategies in sepsis
 - Ran Liu, Joseph L. Greenstein, James C. Fackler, Jules Bergmann, Melania M. Bembea, and Raimond L. Winslow. arXiv, 2021.
 - Improving Long-Term Metrics in Recommendation Systems using Short-Horizon Offline RL
 - Bogdan Mazouze, Paul Mineiro, Pavithra Srinath, Reza Sharifi Sedeh, Doina Precup, and Adith Swaminathan. arXiv, 2021.
 - Safe Model-based Off-policy Reinforcement Learning for Eco-Driving in Connected and Automated Hybrid Electric Vehicles
-

-
- Zhaoxuan Zhu, Nicola Pivaro, Shobhit Gupta, Abhishek Gupta, and Marcello Canova. arXiv, 2021.
 - pH-RL: A personalization architecture to bring reinforcement learning to health practice
 - Ali el Hassouni, Mark Hoogendoorn, Marketa Ciharova, Annet Kleiboer, Khadicha Amarti, Vesa Muhonen, Heleen Riper, and A. E. Eiben. arXiv, 2021.
 - DeepThermal: Combustion Optimization for Thermal Power Generating Units Using Offline Reinforcement Learning [podcast]
 - Xianyuan Zhan, Haoran Xu, Yue Zhang, Yusen Huo, Xiangyu Zhu, Honglei Yin, and Yu Zheng. arXiv, 2021.
 - Personalization for Web-based Services using Offline Reinforcement Learning
 - Pavlos Athanasios Apostolopoulos, Zehui Wang, Hanson Wang, Chad Zhou, Kittipat Virochsiri, Norm Zhou, and Igor L. Markov. arXiv, 2021.
 - BCORLE(λ): An Offline Reinforcement Learning and Evaluation Framework for Coupons Allocation in E-commerce Market
 - Yang Zhang, Bo Tang, Qingyu Yang, Dou An, Hongyin Tang, Chenyang Xi, Xueying Li, and Feiyu Xiong. NeurIPS, 2021.
 - Safe Driving via Expert Guided Policy Optimization [website] [code]
 - Zhenghao Peng, Quanyi Li, Chunxiao Liu, and Bolei Zhou. CoRL, 2021.
 - A General Offline Reinforcement Learning Framework for Interactive Recommendation
 - Teng Xiao and Donglin Wang. AAAI, 2021.
 - Value Function is All You Need: A Unified Learning Framework for Ride Hailing Platforms
 - Xiaocheng Tang, Fan Zhang, Zhiwei (Tony)Qin, Yansheng Wang, Dingyuan Shi, Bingchen Song, Yongxin Tong, Hongtu Zhu, and Jieping Ye. KDD, 2021.
 - Discovering an Aid Policy to Minimize Student Evasion Using Offline Reinforcement Learning
 - Leandro M. de Lima and Renato A. Krohling. IJCNN, 2021.
 - Learning robust driving policies without online exploration
 - Daniel Graves, Nhat M. Nguyen, Kimia Hassanzadeh, Jun Jin, and Jun Luo. ICRA, 2021.
 - Engagement Rewarded Actor-Critic with Conservative Q-Learning for Speech-Driven Laughter Backchannel Generation
-

-
- Öykü Zeynep Bayramoğlu, Engin Erzin, Tevfik Metin Sezgin, and Yücel Yemez. ICMI, 2021.
 - Network Intrusion Detection Based on Extended RBF Neural Network With Offline Reinforcement Learning
 - Manuel Lopez-Martin, Antonio Sanchez-Esguevillas, Juan Ignacio Arribas, and Belen Carro. IEEE Access, 2021.
 - Towards Accelerating Offline RL based Recommender Systems
 - Mayank Mishra, Rekha Singhal, and Ravi Singh. AIMLSystems, 2021.
 - Offline Meta-level Model-based Reinforcement Learning Approach for Cold-Start Recommendation
 - Yanan Wang, Yong Ge, Li Li, Rui Chen, and Tong Xu. arXiv, 2020.
 - Batch-Constrained Distributional Reinforcement Learning for Session-based Recommendation
 - Diksha Garg, Priyanka Gupta, Pankaj Malhotra, Lovekesh Vig, and Gautam Shroff. arXiv, 2020.
 - An Empirical Study of Representation Learning for Reinforcement Learning in Healthcare
 - Taylor W. Killian, Haoran Zhang, Jayakumar Subramanian, Mehdi Fatemi, and Marzyeh Ghassemi. arXiv, 2020.
 - Learning from Human Feedback: Challenges for Real-World Reinforcement Learning in NLP
 - Julia Kreutzer, Stefan Riezler, and Carolin Lawrence. arXiv, 2020.
 - Remote Electrical Tilt Optimization via Safe Reinforcement Learning
 - Filippo Vannella, Grigorios Iakovidis, Ezeddin Al Hakim, Erik Aumayr, and Saman Feghhi. arXiv, 2020.
 - An Optimistic Perspective on Offline Reinforcement Learning [website] [blog]
 - Rishabh Agarwal, Dale Schuurmans, and Mohammad Norouzi. ICML, 2020.
 - Policy Teaching via Environment Poisoning: Training-time Adversarial Attacks against Reinforcement Learning
 - Amin Rakhsha, Goran Radanovic, Rati Devidze, Xiaojin Zhu, and Adish Singla. ICML, 2020.
 - Offline Contextual Multi-armed Bandits for Mobile Health Interventions: A Case Study on Emotion Regulation
 - Mawulolo K. Ameko, Miranda L. Beltzer, Lihua Cai, Mehdi Boukhechba, Bethany A. Teachman, and Laura E. Barnes. RecSys, 2020.
-

-
- Human-centric Dialog Training via Offline Reinforcement Learning
 - Natasha Jaques, Judy Hanwen Shen, Asma Ghandeharioun, Craig Ferguson, Agata Lapedriza, Noah Jones, Shixiang Shane Gu, and Rosalind Picard. EMNLP, 2020.
 - Definition and evaluation of model-free coordination of electrical vehicle charging with reinforcement learning
 - Nasrin Sadeghianpourhamami, Johannes Deleu, and Chris Develder. IEEE T SMART GRID, 2020.
 - Optimal Tap Setting of Voltage Regulation Transformers Using Batch Reinforcement Learning
 - Hanchen Xu, Alejandro D. Domínguez-García, and Peter W. Sauer. IEEE T POWER SYSTEMS, 2020.
 - Way Off-Policy Batch Deep Reinforcement Learning of Implicit Human Preferences in Dialog
 - Natasha Jaques, Asma Ghandeharioun, Judy Hanwen Shen, Craig Ferguson, Agata Lapedriza, Noah Jones, Shixiang Gu, and Rosalind Picard. arXiv, 2019.
 - Optimized cost function for demand response coordination of multiple EV charging stations using reinforcement learning
 - Manu Lahariya, Nasrin Sadeghianpourhamami, and Chris Develder. BuildSys, 2019.
 - A Clustering-Based Reinforcement Learning Approach for Tailored Personalization of E-Health Interventions
 - Ali el Hassouni, Mark Hoogendoorn, Martijn van Otterlo, A. E. Eiben, Vesa Muhonen, and Eduardo Barbaro. arXiv, 2018.
 - Generating Interpretable Fuzzy Controllers using Particle Swarm Optimization and Genetic Programming
 - Daniel Hein, Steffen Udluft, and Thomas A. Runkler. GECCO, 2018.
 - End-to-End Offline Goal-Oriented Dialog Policy Learning via Policy Gradient
 - Li Zhou, Kevin Small, Oleg Rokhlenko, and Charles Elkan. arXiv, 2017.
 - Batch Reinforcement Learning on the Industrial Benchmark: First Experiences
 - Daniel Hein, Steffen Udluft, Michel Tokic, Alexander Hentschel, Thomas A. Runkler, and Volkmar Sterzing. IJCNN, 2017.
 - Policy Networks with Two-Stage Training for Dialogue Systems
 - Mehdi Fatemi, Layla El Asri, Hannes Schulz, Jing He, and Kaheer Suleman. SIGDial, 2016.
-

-
- Adaptive Treatment of Epilepsy via Batch-mode Reinforcement Learning
 - Arthur Guez, Robert D. Vincent, Massimo Avoli, and Joelle Pineau. IAAI, 2008.

Off-Policy Evaluation and Learning: Theory/Methods

Off-Policy Evaluation: Contextual Bandits

- Off-Policy Evaluation of Slate Bandit Policies via Optimizing Abstraction
 - Haruka Kiyohara, Masahiro Nomura, and Yuta Saito. WWW, 2024.
- Distributionally Robust Policy Evaluation under General Covariate Shift in Contextual Bandits
 - Yihong Guo, Hao Liu, Yisong Yue, and Anqi Liu. arXiv, 2024.
- Off-Policy Evaluation for Large Action Spaces via Conjunct Effect Modeling
 - Yuta Saito, Qingyang Ren, and Thorsten Joachims. ICML, 2023.
- Multiply Robust Off-policy Evaluation and Learning under Truncation by Death
 - Jianing Chu, Shu Yang, and Wenbin Lu. ICML, 2023.
- Off-Policy Evaluation of Ranking Policies under Diverse User Behavior
 - Haruka Kiyohara, Masatoshi Uehara, Yusuke Narita, Nobuyuki Shimizu, Yasuo Yamamoto, and Yuta Saito. KDD, 2023.
- Policy-Adaptive Estimator Selection for Off-Policy Evaluation
 - Takuma Udagawa, Haruka Kiyohara, Yusuke Narita, Yuta Saito, and Kei Tateno. AAAI, 2023.
- Variance-Optimal Augmentation Logging for Counterfactual Evaluation in Contextual Bandits
 - Aaron David Tucker and Thorsten Joachims. WSDM, 2023.
- Offline Policy Evaluation in Large Action Spaces via Outcome-Oriented Action Grouping
 - Jie Peng, Hao Zou, Jiashuo Liu, Shaoming Li, Yibao Jiang, Jian Pei, and Peng Cui. WWW, 2023.
- Off-Policy Evaluation for Large Action Spaces via Policy Convolution
 - Noveen Sachdeva, Lequn Wang, Dawen Liang, Nathan Kallus, and Julian McAuley. arXiv, 2023.
- Distributional Off-Policy Evaluation for Slate Recommendations
 - Shreyas Chaudhari, David Arbour, Georgios Theodorou, and Nikos Vlassis. arXiv, 2023.

-
- Debiased Machine Learning and Network Cohesion for Doubly-Robust Differential Reward Models in Contextual Bandits
 - Easton K. Huch, Jieru Shi, Madeline R. Abbott, Jessica R. Golbus, Alexander Moreno, and Walter H. Dempsey. arXiv, 2023.
 - Doubly Robust Estimator for Off-Policy Evaluation with Large Action Spaces
 - Tatsuhiko Shimizu. arXiv, 2023.
 - Offline Policy Evaluation with Out-of-Sample Guarantees
 - Sofia Ek and Dave Zachariah. arXiv, 2023.
 - Quantile Off-Policy Evaluation via Deep Conditional Generative Learning
 - Yang Xu, Chengchun Shi, Shikai Luo, Lan Wang, and Rui Song. arXiv, 2023.
 - Doubly Robust Off-Policy Evaluation for Ranking Policies under the Cascade Behavior Model [code]
 - Haruka Kiyohara, Yuta Saito, Tatsuya Matsuhiko, Yusuke Narita, Nobuyuki Shimizu, and Yasuo Yamamoto. WSDM, 2022.
 - Off-Policy Evaluation for Large Action Spaces via Embeddings [code] [video]
 - Yuta Saito and Thorsten Joachims. ICML, 2022.
 - Doubly Robust Distributionally Robust Off-Policy Evaluation and Learning
 - Nathan Kallus, Xiaojie Mao, Kaiwen Wang, and Zhengyuan Zhou. ICML, 2022.
 - Local Metric Learning for Off-Policy Evaluation in Contextual Bandits with Continuous Actions
 - Haanvid Lee, Jongmin Lee, Yunseon Choi, Wonseok Jeon, Byung-Jun Lee, Yung-Kyun Noh, and Kee-Eung Kim. NeurIPS, 2022.
 - Conformal Off-Policy Prediction in Contextual Bandits
 - Muhammad Faaiz Taufiq, Jean-Francois Ton, Rob Cornish, Yee Whye Teh, and Arnaud Doucet. NeurIPS, 2022.
 - Off-Policy Evaluation with Policy-Dependent Optimization Response
 - Wenshuo Guo, Michael I. Jordan, and Angela Zhou. NeurIPS, 2022.
 - Off-Policy Evaluation with Deficient Support Using Side Information
 - Nicolò Felicioni, Maurizio Ferrari Dacrema, Marcello Restelli, and Paolo Cremonesi. NeurIPS, 2022.
-

-
- Towards Robust Off-Policy Evaluation via Human Inputs
 - Harvineet Singh, Shalmali Joshi, Finale Doshi-Velez, and Himabindu Lakkaraju. AIES, 2022.
 - Off-policy evaluation for learning-to-rank via interpolating the item-position model and the position-based model
 - Alexander Buchholz, Ben London, Giuseppe di Benedetto, and Thorsten Joachims. arXiv, 2022.
 - Bayesian Counterfactual Mean Embeddings and Off-Policy Evaluation
 - Diego Martinez-Taboada and Dino Sejdinovic. arXiv, 2022.
 - Anytime-valid off-policy inference for contextual bandits
 - Ian Waudby-Smith, Lili Wu, Aaditya Ramdas, Nikos Karampatziakis, and Paul Mineiro. arXiv, 2022.
 - Off-policy estimation of linear functionals: Non-asymptotic theory for semi-parametric efficiency
 - Wenlong Mou, Martin J. Wainwright, and Peter L. Bartlett. arXiv, 2022.
 - Off-Policy Evaluation in Embedded Spaces
 - Jaron J. R. Lee, David Arbour, and Georgios Theodorou. arXiv, 2022.
 - Safe Exploration for Efficient Policy Evaluation and Comparison
 - Runzhe Wan, Branislav Kveton, and Rui Song. arXiv, 2022.
 - Inverse Propensity Score based offline estimator for deterministic ranking lists using position bias
 - Nick Wood and Sumit Sidana. arXiv, 2022.
 - Subgaussian and Differentiable Importance Sampling for Off-Policy Evaluation and Learning
 - Alberto Maria Metelli, Alessio Russo, Marcello Restelli. NeurIPS, 2021.
 - Control Variates for Slate Off-Policy Evaluation
 - Nikos Vlassis, Ashok Chandrashekar, Fernando Amat Gil, and Nathan Kallus. NeurIPS, 2021.
 - Deep Jump Learning for Off-Policy Evaluation in Continuous Treatment Settings
 - Hengrui Cai, Chengchun Shi, Rui Song, and Wenbin Lu. NeurIPS, 2021.

-
- Optimal Off-Policy Evaluation from Multiple Logging Policies [code]
 - Nathan Kallus, Yuta Saito, and Masatoshi Uehara. ICML, 2021.
 - Off-policy Confidence Sequences
 - Nikos Karampatziakis, Paul Mineiro, and Aaditya Ramdas. ICML, 2021.
 - Confident Off-Policy Evaluation and Selection through Self-Normalized Importance Weighting [video]
 - Ilja Kuzborskij, Claire Vernade, András György, and Csaba Szepesvári. AISTATS, 2021.
 - Off-Policy Evaluation Using Information Borrowing and Context-Based Switching
 - Sutanoy Dasgupta, Yabo Niu, Kishan Panaganti, Dileep Kalathil, Debdeep Pati, and Bani Mallick. arXiv, 2021.
 - Identification of Subgroups With Similar Benefits in Off-Policy Policy Evaluation
 - Ramtin Keramati, Omer Gottesman, Leo Anthony Celi, Finale Doshi-Velez, and Emma Brunskill. arXiv, 2021.
 - Robust On-Policy Data Collection for Data-Efficient Policy Evaluation
 - Rujie Zhong, Josiah P. Hanna, Lukas Schäfer, and Stefano V. Albrecht. arXiv, 2021.
 - Off-Policy Evaluation via Adaptive Weighting with Data from Contextual Bandits
 - Ruohan Zhan, Vitor Hadad, David A. Hirshberg, and Susan Athey. arXiv, 2021.
 - Off-Policy Risk Assessment in Contextual Bandits
 - Audrey Huang, Liu Leqi, Zachary C. Lipton, and Kamyar Azizzadenesheli. arXiv, 2021.
 - Off-Policy Evaluation of Slate Policies under Bayes Risk
 - Nikos Vlassis, Fernando Amat Gil, and Ashok Chandrashekar. arXiv, 2021.
 - A Practical Guide of Off-Policy Evaluation for Bandit Problems
 - Masahiro Kato, Kenshi Abe, Kaito Ariu, and Shota Yasui. arXiv, 2020.
 - Off-Policy Evaluation and Learning for External Validity under a Covariate Shift
 - Masatoshi Uehara, Masahiro Kato, and Shota Yasui. NeurIPS, 2020.
 - Counterfactual Evaluation of Slate Recommendations with Sequential Reward Interactions
 - James McInerney, Brian Brost, Praveen Chandar, Rishabh Mehrotra, and Ben Carterette. KDD, 2020.

-
- Doubly robust off-policy evaluation with shrinkage
 - Yi Su, Maria Dimakopoulou, Akshay Krishnamurthy, and Miroslav Dudik. ICML, 2020.
 - Adaptive Estimator Selection for Off-Policy Evaluation [video]
 - Yi Su, Pavithra Srinath, and Akshay Krishnamurthy. ICML, 2020.
 - Distributionally Robust Policy Evaluation and Learning in Offline Contextual Bandits
 - Nian Si, Fan Zhang, Zhengyuan Zhou, and Jose Blanchet. ICML, 2020.
 - Improving Offline Contextual Bandits with Distributional Robustness
 - Otmane Sakhi, Louis Faury, and Flavian Vasile. arXiv, 2020.
 - Balanced Off-Policy Evaluation in General Action Spaces
 - Arjun Sondhi, David Arbour, and Drew Dimmery. AISTATS, 2019.
 - Policy Evaluation with Latent Confounders via Optimal Balance
 - Andrew Bennett and Nathan Kallus. NeuIPS, 2019.
 - On the Design of Estimators for Bandit Off-Policy Evaluation
 - Nikos Vlassis, Aurelien Bibaut, Maria Dimakopoulou, and Tony Jebara. ICML, 2019.
 - CAB: Continuous Adaptive Blending for Policy Evaluation and Learning
 - Yi Su, Lequn Wang, Michele Santacatterina, and Thorsten Joachims. ICML, 2019.
 - Focused Context Balancing for Robust Offline Policy Evaluation
 - Hao Zou, Kun Kuang, Boqi Chen, Peixuan Chen, and Peng Cui. KDD, 2019.
 - When People Change their Mind: Off-Policy Evaluation in Non-Stationary Recommendation Environments
 - Rolf Jagerman, Ilya Markov, and Maarten de Rijke. WSDM, 2019.
 - Policy Evaluation and Optimization with Continuous Treatments
 - Nathan Kallus and Angela Zhou. AISTATS, 2019.
 - Confounding-Robust Policy Improvement
 - Nathan Kallus and Angela Zhou. NeuIPS, 2018.
 - Balanced Policy Evaluation and Learning
 - Nathan Kallus. NeuIPS, 2018.
-

-
- Offline Evaluation of Ranking Policies with Click Models
 - Shuai Li, Yasin Abbasi-Yadkori, Branislav Kveton, S. Muthukrishnan, Vishwa Vinay, and Zheng Wen. KDD, 2018.
 - Effective Evaluation using Logged Bandit Feedback from Multiple Loggers
 - Aman Agarwal, Soumya Basu, Tobias Schnabel, and Thorsten Joachims. KDD, 2018.
 - Off-policy Evaluation for Slate Recommendation
 - Adith Swaminathan, Akshay Krishnamurthy, Alekh Agarwal, Miroslav Dudík, John Langford, Damien Jose, and Imed Zitouni. NeurIPS, 2017.
 - Optimal and Adaptive Off-policy Evaluation in Contextual Bandits
 - Yu-Xiang Wang, Alekh Agarwal, and Miroslav Dudik. ICML, 2017.
 - Data-Efficient Policy Evaluation Through Behavior Policy Search
 - Josiah P. Hanna, Philip S. Thomas, Peter Stone, and Scott Niekum. ICML, 2017.
 - Doubly Robust Policy Evaluation and Optimization
 - Miroslav Dudík, Dumitru Erhan, John Langford, and Lihong Li. ICML, 2011.
 - Unbiased Offline Evaluation of Contextual-bandit-based News Article Recommendation Algorithms
 - Lihong Li, Wei Chu, John Langford, and Xuanhui Wang. WSDM, 2011.

Off-Policy Evaluation: Reinforcement Learning

- Distributional Off-policy Evaluation with Bellman Residual Minimization
 - Sungeel Hong, Zhengling Qi, and Raymond K. W. Wong. arXiv, 2024.
- Future-Dependent Value-Based Off-Policy Evaluation in POMDPs
 - Masatoshi Uehara, Haruka Kiyohara, Andrew Bennett, Victor Chernozhukov, Nan Jiang, Nathan Kallus, Chengchun Shi, and Wen Sun. NeurIPS, 2023.
- Marginal Density Ratio for Off-Policy Evaluation in Contextual Bandits
 - Muhammad Faaiz Taufiq, Arnaud Doucet, Rob Cornish, and Jean-Francois Ton. NeurIPS, 2023.
- State-Action Similarity-Based Representations for Off-Policy Evaluation

-
- Brahma S. Pavse and Josiah P. Hanna. NeurIPS, 2023.
 - Off-Policy Evaluation for Human Feedback
 - Qitong Gao, Juncheng Dong, Vahid Tarokh, Min Chi, and Miroslav Pajic. NeurIPS, 2023.
 - Counterfactual-Augmented Importance Sampling for Semi-Offline Policy Evaluation
 - Shengpu Tang and Jenna Wiens. NeurIPS, 2023.
 - An Instrumental Variable Approach to Confounded Off-Policy Evaluation
 - Yang Xu, Jin Zhu, Chengchun Shi, Shikai Luo, and Rui Song. ICML, 2023.
 - Semiparametrically Efficient Off-Policy Evaluation in Linear Markov Decision Processes
 - Chuhan Xie, Wenhao Yang, and Zhihua Zhang. ICML, 2023.
 - Distributional Offline Policy Evaluation with Predictive Error Guarantees
 - Runzhe Wu, Masatoshi Uehara, and Wen Sun. ICML, 2023.
 - The Optimal Approximation Factors in Misspecified Off-Policy Value Function Estimation
 - Philip Amortila, Nan Jiang, and Csaba Szepesvári. ICML, 2023.
 - Revisiting Bellman Errors for Offline Model Selection [code]
 - Joshua P. Zitovsky, Daniel de Marchi, Rishabh Agarwal, and Michael R. Kosorok. ICML, 2023.
 - Scaling Marginalized Importance Sampling to High-Dimensional State-Spaces via State Abstraction
 - Brahma S. Pavse and Josiah P. Hanna. AAAI, 2023.
 - Variational Latent Branching Model for Off-Policy Evaluation
 - Qitong Gao, Ge Gao, Min Chi, and Miroslav Pajic. ICLR, 2023.
 - Multiple-policy High-confidence Policy Evaluation
 - Chris Dann, Mohammad Ghavamzadeh, and Teodor V. Marinov. AISTATS, 2023.
 - Off-Policy Evaluation with Online Adaptation for Robot Exploration in Challenging Environments
 - Yafei Hu, Junyi Geng, Chen Wang, John Keller, and Sebastian Scherer. RA-L, 2023.
 - Conservative Exploration for Policy Optimization via Off-Policy Policy Evaluation

-
- Paul Daoudi, Mathias Formoso, Othman Gaizi, Achraf Azize, and Evrard Garcelon. arXiv, 2023.
 - Robust Offline Policy Evaluation and Optimization with Heavy-Tailed Rewards
 - Jin Zhu, Runzhe Wan, Zhengling Qi, Shikai Luo, and Chengchun Shi. arXiv, 2023.
 - When is Offline Policy Selection Sample Efficient for Reinforcement Learning?
 - Vincent Liu, Prabhat Nagarajan, Andrew Patterson, and Martha White. arXiv, 2023.
 - Sample Complexity of Preference-Based Nonparametric Off-Policy Evaluation with Deep Networks
 - Zihao Li, Xiang Ji, Minshuo Chen, and Mengdi Wang. arXiv, 2023.
 - Evaluation of Active Feature Acquisition Methods for Static Feature Settings
 - Henrik von Kleist, Alireza Zamanian, Ilya Shpitser, and Narges Ahmidi. arXiv, 2023.
 - Distributional Shift-Aware Off-Policy Interval Estimation: A Unified Error Quantification Framework
 - Wenzhuo Zhou, Yuhan Li, Ruqing Zhu, and Annie Qu. arXiv, 2023.
 - Marginalized Importance Sampling for Off-Environment Policy Evaluation
 - Pulkit Katdare, Nan Jiang, and Katherine Driggs-Campbell. arXiv, 2023.
 - Statistically Efficient Variance Reduction with Double Policy Estimation for Off-Policy Evaluation in Sequence-Modeled Reinforcement Learning
 - Hanhan Zhou, Tian Lan, and Vaneet Aggarwal. arXiv, 2023.
 - Asymptotically Unbiased Off-Policy Policy Evaluation when Reusing Old Data in Nonstationary Environments
 - Vincent Liu, Yash Chandak, Philip Thomas, and Martha White. arXiv, 2023.
 - Off-policy Evaluation in Doubly Inhomogeneous Environments
 - Zeyu Bian, Chengchun Shi, Zhengling Qi, and Lan Wang. arXiv, 2023.
 - Offline Policy Evaluation for Reinforcement Learning with Adaptively Collected Data
 - Sunil Madhow, Dan Xiao, Ming Yin, and Yu-Xiang Wang. arXiv, 2023.
 - $\pi_2\text{vec}$: Policy Representations with Successor Features
 - Gianluca Scarpellini, Ksenia Konyushkova, Claudio Fantacci, Tom Le Paine, Yutian Chen, and Misha Denil. arXiv, 2023.
-

-
- Conformal Off-Policy Evaluation in Markov Decision Processes
 - Daniele Foffano, Alessio Russo, and Alexandre Proutiere. arXiv, 2023.
 - Hallucinated Adversarial Control for Conservative Offline Policy Evaluation
 - Jonas Rothfuss, Bhavya Sukhija, Tobias Birchler, Parnian Kassraie, and Andreas Krause. arXiv, 2023.
 - Robust Fitted-Q-Evaluation and Iteration under Sequentially Exogenous Unobserved Confounders
 - David Bruns-Smith and Angela Zhou. arXiv, 2023.
 - Minimax Weight Learning for Absorbing MDPs
 - Fengyin Li, Yuqiang Li, and Xianyi Wu. arXiv, 2023.
 - Improving Monte Carlo Evaluation with Offline Data
 - Shuze Liu and Shangtong Zhang. arXiv, 2023.
 - First-order Policy Optimization for Robust Policy Evaluation
 - Yan Li and Guanghui Lan. arXiv, 2023.
 - A Minimax Learning Approach to Off-Policy Evaluation in Confounded Partially Observable Markov Decision Processes
 - Chengchun Shi, Masatoshi Uehara, Jiawei Huang, and Nan Jiang. ICML, 2022.
 - On Well-posedness and Minimax Optimal Rates of Nonparametric Q-function Estimation in Off-policy Evaluation
 - Xiaohong Chen and Zhengling Qi. ICML, 2022.
 - Learning Bellman Complete Representations for Offline Policy Evaluation
 - Jonathan Chang, Kaiwen Wang, Nathan Kallus, and Wen Sun. ICML, 2022.
 - Supervised Off-Policy Ranking
 - Yue Jin, Yue Zhang, Tao Qin, Xudong Zhang, Jian Yuan, Houqiang Li, and Tie-Yan Liu. ICML, 2022.
 - Off-Policy Fitted Q-Evaluation with Differentiable Function Approximators: Z-Estimation and Inference Theory
 - Ruiqi Zhang, Xuezhou Zhang, Chengzhuo Ni, and Mengdi Wang. ICML, 2022.

-
- Beyond the Return: Off-policy Function Estimation under User-specified Error-measuring Distributions
 - Audrey Huang and Nan Jiang. NeurIPS, 2022.
 - Oracle Inequalities for Model Selection in Offline Reinforcement Learning
 - Jonathan N. Lee, George Tucker, Ofir Nachum, Bo Dai, and Emma Brunskill. NeurIPS, 2022.
 - Off-Policy Evaluation for Episodic Partially Observable Markov Decision Processes under Non-Parametric Models
 - Rui Miao, Zhengling Qi, and Xiaoke Zhang. NeurIPS, 2022.
 - Off-Policy Evaluation for Action-Dependent Non-stationary Environments
 - Yash Chandak, Shiv Shankar, Nathaniel D. Bastian, Bruno Castro da Silva, Emma Brunskill, and Philip S. Thomas. NeurIPS, 2022.
 - Stateful Offline Contextual Policy Evaluation and Learning
 - Nathan Kallus, and Angela Zhou. AISTATS, 2022.
 - Off-Policy Risk Assessment for Markov Decision Processes
 - Audrey Huang, Liu Leqi, Zachary Lipton, and Kamyar Azizzadenesheli. AISTATS, 2022.
 - Offline Reinforcement Learning for Human-Guided Human-Machine Interaction with Private Information
 - Zuyue Fu, Zhengling Qi, Zhuoran Yang, Zhaoran Wang, and Lan Wang. arXiv, 2022.
 - Offline Policy Evaluation and Optimization under Confounding
 - Kevin Tan, Yangyi Lu, Chinmaya Kausik, Yixin Wang, and Ambuj Tewari. arXiv, 2022.
 - Bridging the Gap Between Offline and Online Reinforcement Learning Evaluation Methodologies
 - Shivakanth Sujit, Pedro H. M. Braga, Jorg Bornschein, and Samira Ebrahimi Kahou. arXiv, 2022.
 - Safe Evaluation For Offline Learning: Are We Ready To Deploy?
 - Hager Radi, Josiah P. Hanna, Peter Stone, and Matthew E. Taylor. arXiv, 2022.
 - Low Variance Off-policy Evaluation with State-based Importance Sampling
 - David M. Bossens and Philip Thomas. arXiv, 2022.

-
- Statistical Estimation of Confounded Linear MDPs: An Instrumental Variable Approach
 - Miao Lu, Wenhao Yang, Liangyu Zhang, and Zhihua Zhang. arXiv, 2022.
 - Offline Estimation of Controlled Markov Chains: Minimax Nonparametric Estimators and Sample Efficiency
 - Imon Banerjee, Harsha Honnappa, and Vinayak Rao. arXiv, 2022.
 - Sample Complexity of Nonparametric Off-Policy Evaluation on Low-Dimensional Manifolds using Deep Networks
 - Xiang Ji, Minshuo Chen, Mengdi Wang, and Tuo Zhao. arXiv, 2022.
 - A Sharp Characterization of Linear Estimators for Offline Policy Evaluation
 - Juan C. Perdomo, Akshay Krishnamurthy, Peter Bartlett, and Sham Kakade. arXiv, 2022.
 - A Multi-Agent Reinforcement Learning Framework for Off-Policy Evaluation in Two-sided Markets [code]
 - Chengchun Shi, Runzhe Wan, Ge Song, Shikai Luo, Rui Song, and Hongtu Zhu. arXiv, 2022.
 - A Theoretical Framework of Almost Hyperparameter-free Hyperparameter Selection Methods for Offline Policy Evaluation
 - Kohei Miyaguchi. arXiv, 2022.
 - SOPE: Spectrum of Off-Policy Estimators
 - Christina J. Yuan, Yash Chandak, Stephen Giguere, Philip S. Thomas, and Scott Niekum. NeurIPS, 2021.
 - Unifying Gradient Estimators for Meta-Reinforcement Learning via Off-Policy Evaluation
 - Yunhao Tang, Tadashi Kozuno, Mark Rowland, Rémi Munos, and Michal Valko. NeurIPS, 2021.
 - Variance-Aware Off-Policy Evaluation with Linear Function Approximation
 - Yifei Min, Tianhao Wang, Dongruo Zhou, and Quanquan Gu. NeurIPS, 2021.
 - Universal Off-Policy Evaluation
 - Yash Chandak, Scott Niekum, Bruno Castro da Silva, Erik Learned-Miller, Emma Brunskill, and Philip S. Thomas. NeurIPS, 2021.
 - Towards Hyperparameter-free Policy Selection for Offline Reinforcement Learning
 - Siyuan Zhang and Nan Jiang. NeurIPS, 2021.

-
- Optimal Uniform OPE and Model-based Offline Reinforcement Learning in Time-Homogeneous, Reward-Free and Task-Agnostic Settings
 - Ming Yin and Yu-Xiang Wang. NeurIPS, 2021.
 - State Relevance for Off-Policy Evaluation
 - Simon P. Shen, Yecheng Jason Ma, Omer Gottesman, and Finale Doshi-Velez. ICML, 2021.
 - Bootstrapping Fitted Q-Evaluation for Off-Policy Inference
 - Botao Hao, Xiang Ji, Yaqi Duan, Hao Lu, Csaba Szepesvari, and Mengdi Wang. ICML, 2021.
 - Deeply-Debiased Off-Policy Interval Estimation
 - Chengchun Shi, Runzhe Wan, Victor Chernozhukov, and Rui Song. ICML, 2021.
 - Autoregressive Dynamics Models for Offline Policy Evaluation and Optimization
 - Michael R. Zhang, Tom Le Paine, Ofir Nachum, Cosmin Paduraru, George Tucker, Ziyu Wang, Mohammad Norouzi. ICLR, 2021.
 - Minimax Model Learning
 - Cameron Voloshin, Nan Jiang, and Yisong Yue. AISTATS, 2021.
 - Off-policy Evaluation in Infinite-Horizon Reinforcement Learning with Latent Confounders
 - Andrew Bennett, Nathan Kallus, Lihong Li, and Ali Mousavi. AISTATS, 2021.
 - High-Confidence Off-Policy (or Counterfactual) Variance Estimation
 - Yash Chandak, Shiv Shankar, and Philip S. Thomas. AAAI, 2021.
 - Debiased Off-Policy Evaluation for Recommendation Systems
 - Yusuke Narita, Shota Yasui, and Kohei Yata. RecSys, 2021.
 - Pessimistic Model Selection for Offline Deep Reinforcement Learning
 - Chao-Han Huck Yang, Zhengling Qi, Yifan Cui, and Pin-Yu Chen. arXiv, 2021.
 - Proximal Reinforcement Learning: Efficient Off-Policy Evaluation in Partially Observed Markov Decision Processes
 - Andrew Bennett and Nathan Kallus. arXiv, 2021.
 - Off-Policy Evaluation in Partially Observed Markov Decision Processes
 - Yuchen Hu and Stefan Wager. arXiv, 2021.

-
- A Spectral Approach to Off-Policy Evaluation for POMDPs
 - Yash Nair and Nan Jiang. arXiv, 2021.
 - Projected State-action Balancing Weights for Offline Reinforcement Learnings
 - Jiayi Wang, Zhengling Qi, and Raymond K.W. Wong. arXiv, 2021.
 - Active Offline Policy Selection
 - Ksenia Konyushkova, Yutian Chen, Thomas Paine, Caglar Gulcehre, Cosmin Paduraru, Daniel J Mankowitz, Misha Denil, and Nando de Freitas. arXiv, 2021.
 - On Instrumental Variable Regression for Deep Offline Policy Evaluation
 - Yutian Chen, Liyuan Xu, Caglar Gulcehre, Tom Le Paine, Arthur Gretton, Nando de Freitas, and Arnaud Doucet. arXiv, 2021.
 - Average-Reward Off-Policy Policy Evaluation with Function Approximation
 - Shangdong Zhang, Yi Wan, Richard S. Sutton, and Shimon Whiteson. arXiv, 2021.
 - Sequential causal inference in a single world of connected units
 - Aurelien Bibaut, Maya Petersen, Nikos Vlassis, Maria Dimakopoulou, and Mark van der Laan, arXiv, 2021.
 - Off-policy Policy Evaluation For Sequential Decisions Under Unobserved Confounding
 - Hongseok Namkoong, Ramtin Keramati, Steve Yadlowsky, and Emma Brunskill. NeurIPS, 2020.
 - CoinDICE: Off-Policy Confidence Interval Estimation
 - Bo Dai, Ofir Nachum, Yinlam Chow, Lihong Li, Csaba Szepesvari, and Dale Schuurmans. NeurIPS, 2020.
 - Off-Policy Interval Estimation with Lipschitz Value Iteration
 - Ziyang Tang, Yihao Feng, Na Zhang, Jian Peng, and Qiang Liu. NeurIPS, 2020.
 - Off-Policy Evaluation via the Regularized Lagrangian
 - Mengjiao Yang, Ofir Nachum, Bo Dai, Lihong Li, and Dale Schuurmans. NeurIPS, 2020.
 - Minimax Value Interval for Off-Policy Evaluation and Policy Optimization
 - Nan Jiang and Jiawei Huang. NeurIPS, 2020.
 - GenDICE: Generalized Offline Estimation of Stationary Values

-
- Ruiyi Zhang, Bo Dai, Lihong Li, and Dale Schuurmans. ICLR, 2020.
 - Infinite-horizon Off-Policy Policy Evaluation with Multiple Behavior Policies
 - Xinyun Chen, Lu Wang, Yizhe Hang, Heng Ge, and Hongyuan Zha. ICLR, 2020.
 - Doubly Robust Bias Reduction in Infinite Horizon Off-Policy Estimation
 - Ziyang Tang, Yihao Feng, Lihong Li, Dengyong Zhou, and Qiang Liu. ICLR, 2020.
 - Black-box Off-policy Estimation for Infinite-Horizon Reinforcement Learning
 - Ali Mousavi, Lihong Li, Qiang Liu, and Denny Zhou. ICLR, 2020.
 - GradientDICE: Rethinking Generalized Offline Estimation of Stationary Values
 - Shangdong Zhang, Bo Liu, and Shimon Whiteson. ICML, 2020.
 - Minimax-Optimal Off-Policy Evaluation with Linear Function Approximation
 - Yaqi Duan, Zeyu Jia, and Mengdi Wang. ICML, 2020.
 - Interpretable Off-Policy Evaluation in Reinforcement Learning by Highlighting Influential Transitions
 - Omer Gottesman, Joseph Futoma, Yao Liu, Sonali Parbhoo, Leo Celi, Emma Brunskill, and Finale Doshi-Velez. ICML, 2020.
 - Double Reinforcement Learning for Efficient and Robust Off-Policy Evaluation
 - Nathan Kallus and Masatoshi Uehara. ICML, 2020.
 - Understanding the Curse of Horizon in Off-Policy Evaluation via Conditional Importance Sampling
 - Yao Liu, Pierre-Luc Bacon, and Emma Brunskill. ICML, 2020.
 - Minimax Weight and Q-Function Learning for Off-Policy Evaluation
 - Masatoshi Uehara, Jiawei Huang, and Nan Jiang. ICML, 2020.
 - Accountable Off-Policy Evaluation With Kernel Bellman Statistics
 - Yihao Feng, Tongzheng Ren, Ziyang Tang, and Qiang Liu. ICML, 2020.
 - Asymptotically Efficient Off-Policy Evaluation for Tabular Reinforcement Learning
 - Ming Yin and Yu-Xiang Wang. ICML, 2020.
 - Batch Stationary Distribution Estimation
-

-
- Junfeng Wen, Bo Dai, Lihong Li, and Dale Schuurmans. ICML, 2020.
 - Towards Off-policy Evaluation as a Prerequisite for Real-world Reinforcement Learning in Building Control [video]
 - Bingqing Chen, Ming Jin, Zhe Wang, Tianzhen Hong, and Mario Bergés, RLEM, 2020.
 - Defining Admissible Rewards for High Confidence Policy Evaluation in Batch Reinforcement Learning
 - Niranjani Prasad, Barbara E Engelhardt, and Finale Doshi-Velez. CHIL, 2020.
 - Offline Policy Selection under Uncertainty
 - Mengjiao Yang, Bo Dai, Ofir Nachum, George Tucker, and Dale Schuurmans. arXiv, 2020.
 - Near-Optimal Provable Uniform Convergence in Offline Policy Evaluation for Reinforcement Learning
 - Ming Yin, Yu Bai, and Yu-Xiang Wang. arXiv, 2020.
 - Optimal Mixture Weights for Off-Policy Evaluation with Multiple Behavior Policies
 - Jinlin Lai, Lixin Zou, and Jiaying Song. arXiv, 2020.
 - Kernel Methods for Policy Evaluation: Treatment Effects, Mediation Analysis, and Off-Policy Planning
 - Rahul Singh, Liyuan Xu, and Arthur Gretton. arXiv, 2020.
 - Statistical Bootstrapping for Uncertainty Estimation in Off-Policy Evaluation
 - Ilya Kostrikov and Ofir Nachum. arXiv, 2020.
 - Efficiently Breaking the Curse of Horizon in Off-Policy Evaluation with Double Reinforcement Learning
 - Nathan Kallus and Masatoshi Uehara. arXiv, 2019.
 - Off-Policy Evaluation in Partially Observable Environments
 - Guy Tennenholtz, Uri Shalit, and Shie Mannor. AAAI, 2019.
 - Intrinsically Efficient, Stable, and Bounded Off-Policy Evaluation for Reinforcement Learning
 - Nathan Kallus and Masatoshi Uehara. NeurIPS, 2019.
 - Towards Optimal Off-Policy Evaluation for Reinforcement Learning with Marginalized Importance Sampling

-
- Tengyang Xie, Yifei Ma, and Yu-Xiang Wang. NeuIPS, 2019.
 - Off-Policy Evaluation via Off-Policy Classification
 - Alexander Irpan, Kanishka Rao, Konstantinos Bousmalis, Chris Harris, Julian Ibarz, and Sergey Levine. NeuIPS, 2019.
 - DualDICE: Behavior-Agnostic Estimation of Discounted Stationary Distribution Corrections [software]
 - Ofir Nachum, Yinlam Chow, Bo Dai, Lihong Li. NeurIPS, 2019.
 - Off-Policy Evaluation and Learning from Logged Bandit Feedback: Error Reduction via Surrogate Policy
 - Yuan Xie, Boyi Liu, Qiang Liu, Zhaoran Wang, Yuan Zhou, and Jian Peng. ICLR, 2019.
 - Batch Policy Learning under Constraints [code] [website]
 - Hoang M. Le, Cameron Voloshin, and Yisong Yue. ICML, 2019.
 - More Efficient Off-Policy Evaluation through Regularized Targeted Learning
 - Aurelien Bibaut, Ivana Malenica, Nikos Vlassis, and Mark Van Der Laan. ICML, 2019.
 - Combining parametric and nonparametric models for off-policy evaluation
 - Omer Gottesman, Yao Liu, Scott Sussex, Emma Brunskill, and Finale Doshi-Velez. ICML, 2019.
 - Counterfactual Off-Policy Evaluation with Gumbel-Max Structural Causal Models
 - Michael Oberst and David Sontag. ICML, 2019.
 - Importance Sampling Policy Evaluation with an Estimated Behavior Policy
 - Josiah Hanna, Scott Niekum, and Peter Stone. ICML, 2019.
 - Representation Balancing MDPs for Off-policy Policy Evaluation
 - Yao Liu, Omer Gottesman, Aniruddh Raghu, Matthieu Komorowski, Aldo A. Faisal, Finale Doshi-Velez, and Emma Brunskill. NeuIPS, 2018.
 - Breaking the Curse of Horizon: Infinite-Horizon Off-Policy Estimation
 - Qiang Liu, Lihong Li, Ziyang Tang, and Dengyong Zhou. NeuIPS, 2018.
 - More Robust Doubly Robust Off-policy Evaluation
 - Mehrdad Farajtabar, Yinlam Chow, and Mohammad Ghavamzadeh. ICML, 2018.

-
- Importance Sampling for Fair Policy Selection
 - Shayan Doroudi, Philip Thomas, and Emma Brunskill. UAI, 2017.
 - Predictive Off-Policy Policy Evaluation for Nonstationary Decision Problems, with Applications to Digital Marketing
 - Philip S. Thomas, Georgios Theodorou, Mohammad Ghavamzadeh, Ishan Durugkar, and Emma Brunskill. AAAI, 2017.
 - Consistent On-Line Off-Policy Evaluation
 - Assaf Hallak and Shie Mannor. ICML, 2017.
 - Bootstrapping with Models: Confidence Intervals for Off-Policy Evaluation
 - Josiah P. Hanna, Peter Stone, and Scott Niekum. AAAMS, 2016.
 - Doubly Robust Off-policy Value Evaluation for Reinforcement Learning
 - Nan Jiang and Lihong Li. ICML, 2016.
 - Data-Efficient Off-Policy Policy Evaluation for Reinforcement Learning
 - Philip Thomas and Emma Brunskill. ICML, 2016.
 - High Confidence Policy Improvement
 - Philip Thomas, Georgios Theodorou, and Mohammad Ghavamzadeh. ICML, 2015.
 - High Confidence Off-Policy Evaluation
 - Philip S. Thomas, Georgios Theodorou, and Mohammad Ghavamzadeh. AAAI, 2015.
 - Eligibility Traces for Off-Policy Policy Evaluation
 - Doina Precup, Richard S. Sutton, and Satinder P. Singh. ICML, 2000.

Off-Policy Learning

- Sequential Counterfactual Risk Minimization
 - Houssam Zenati, Eustache Diemert, Matthieu Martin, Julien Mairal, and Pierre Gaillard. ICML, 2023.
- Trajectory-Aware Eligibility Traces for Off-Policy Reinforcement Learning
 - Brett Daley, Martha White, Christopher Amato, and Marlos C. Machado. ICML, 2023.
- Multi-Task Off-Policy Learning from Bandit Feedback

-
- Joey Hong, Branislav Kveton, Sumeet Katariya, Manzil Zaheer, and Mohammad Ghavamzadeh. ICML, 2023.
 - Exponential Smoothing for Off-Policy Learning
 - Imad Aouali, Victor-Emmanuel Brunel, David Rohde, and Anna Korba. ICML, 2023.
 - Counterfactual Learning with General Data-generating Policies
 - Yusuke Narita, Kyohei Okumura, Akihiro Shimizu, and Kohei Yata. AAAI, 2023.
 - Distributionally Robust Policy Gradient for Offline Contextual Bandits
 - Zhouhao Yang, Yihong Guo, Pan Xu, Anqi Liu, and Animashree Anandkumar. AISTATS, 2023.
 - Oracle-Efficient Pessimism: Offline Policy Optimization in Contextual Bandits
 - Lequn Wang, Akshay Krishnamurthy, and Aleksandrs Slivkins. arXiv, 2023.
 - Pessimistic Off-Policy Multi-Objective Optimization
 - Shima Alizadeh, Aniruddha Bhargava, Karthick Gopalswamy, Lalit Jain, Branislav Kveton, and Ge Liu. arXiv, 2023.
 - Unified Off-Policy Learning to Rank: a Reinforcement Learning Perspective
 - Zeyu Zhang, Yi Su, Hui Yuan, Yiran Wu, Rishab Balasubramanian, Qingyun Wu, Huazheng Wang, and Mengdi Wang. arXiv, 2023.
 - Uncertainty-Aware Off-Policy Learning
 - Xiaoying Zhang, Junpu Chen, Hongning Wang, Hong Xie, and Hang Li. arXiv, 2023.
 - Fair Off-Policy Learning from Observational Data
 - Dennis Frauen, Valentyn Melnychuk, and Stefan Feuerriegel. arXiv, 2023.
 - Interpretable Off-Policy Learning via Hyperbox Search
 - Daniel Tschernutter, Tobias Hatt, and Stefan Feuerriegel. ICML, 2022.
 - Offline Policy Optimization with Eligible Actions
 - Yao Liu, Yannis Flet-Berliac, and Emma Brunskill. UAI, 2022.
 - Towards Robust Off-policy Learning for Runtime Uncertainty
 - Da Xu, Yuting Ye, Chuanwei Ruan, and Bo Yang. AAAI, 2022.
 - Safe Optimal Design with Applications in Off-Policy Learning
-

-
- Ruihao Zhu and Branislav Kveton. AISTATS, 2022.
 - Off-Policy Actor-critic for Recommender Systems
 - Minmin Chen, Can Xu, Vince Gatto, Devanshu Jain, Aviral Kumar, and Ed Chi. RecSys, 2022.
 - MGPolicy: Meta Graph Enhanced Off-policy Learning for Recommendations
 - Xiangmeng Wang, Qian Li, Dianer Yu, Zhichao Wang, Hongxu Chen, and Guandong Xu. SIGIR, 2022.
 - Distributionally Robust Policy Learning with Wasserstein Distance
 - Daido Kido. arXiv, 2022.
 - Local Policy Improvement for Recommender Systems
 - Dawen Liang and Nikos Vlassis. arXiv, 2022.
 - Policy learning “without” overlap: Pessimism and generalized empirical Bernstein’s inequality
 - Ying Jin, Zhimei Ren, Zhuoran Yang, and Zhaoran Wang. arXiv, 2022.
 - Fast Offline Policy Optimization for Large Scale Recommendation
 - Otmane Sakhi, David Rohde, and Alexandre Gilotte. arXiv, 2022.
 - Practical Counterfactual Policy Learning for Top-K Recommendations
 - Yaxu Liu, Jui-Nan Yen, Bowen Yuan, Rundong Shi, Peng Yan, and Chih-Jen Lin. KDD, 2022.
 - Boosted Off-Policy Learning
 - Ben London, Levi Lu, Ted Sandler, and Thorsten Joachims. arXiv, 2022.
 - Semi-Counterfactual Risk Minimization Via Neural Networks
 - Gholamali Aminian, Roberto Vega, Omar Rivasplata, Laura Toni, and Miguel Rodrigues. arXiv, 2022.
 - IMO³: Interactive Multi-Objective Off-Policy Optimization
 - Nan Wang, Hongning Wang, Maryam Karimzadehgan, Branislav Kveton, and Craig Boutilier. arXiv, 2022.
 - Pessimistic Off-Policy Optimization for Learning to Rank
 - Matej Cief, Branislav Kveton, and Michal Kompan. arXiv, 2022.
 - Non-Stationary Off-Policy Optimization
-

-
- Joey Hong, Branislav Kveton, Manzil Zaheer, Yinlam Chow, and Amr Ahmed. AISTATS, 2021.
 - Learning from eXtreme Bandit Feedback
 - Romain Lopez, Inderjit Dhillon, and Michael I. Jordan. AAAI, 2021.
 - Generalizing Off-Policy Learning under Sample Selection Bias
 - Tobias Hatt, Daniel Tschernutter, and Stefan Feuerriegel. arXiv, 2021.
 - Conservative Policy Construction Using Variational Autoencoders for Logged Data with Missing Values
 - Mahed Abroshan, Kai Hou Yip, Cem Tekin, and Mihaela van der Schaar. arXiv, 2021.
 - Doubly Robust Off-Policy Value and Gradient Estimation for Deterministic Policies
 - Nathan Kallus and Masatoshi Uehara. NeurIPS, 2020.
 - From Importance Sampling to Doubly Robust Policy Gradient
 - Jiawei Huang and Nan Jiang. ICML, 2020.
 - Efficient Policy Learning from Surrogate-Loss Classification Reductions [code]
 - Andrew Bennett and Nathan Kallus. ICML, 2020.
 - Off-policy Bandits with Deficient Support
 - Noveen Sachdeva, Yi Su, and Thorsten Joachims. KDD, 2020.
 - Off-policy Learning in Two-stage Recommender Systems
 - Jiaqi Ma, Zhe Zhao, Xinyang Yi, Ji Yang, Minmin Chen, Jiaxi Tang, Lichan Hong, and Ed H Chi. WWW, 2020.
 - More Efficient Policy Learning via Optimal Retargeting
 - Nathan Kallus. JASA, 2020.
 - Learning When-to-Treat Policies
 - Xinkun Nie, Emma Brunskill, and Stefan Wager. JASA, 2020.
 - Doubly Robust Off-Policy Learning on Low-Dimensional Manifolds by Deep Neural Networks
 - Minshuo Chen, Hao Liu, Wenjing Liao, and Tuo Zhao. arXiv, 2020.
 - Bandit Overfitting in Offline Policy Learning
 - David Brandfonbrener, William F. Whitney, Rajesh Ranganath, and Joan Bruna. arXiv, 2020.
-

-
- Counterfactual Learning of Continuous Stochastic Policies
 - Houssam Zenati, Alberto Bietti, Matthieu Martin, Eustache Diemert, and Julien Mairal. arXiv, 2020.
 - Top-K Off-Policy Correction for a REINFORCE Recommender System
 - Minmin Chen, Alex Beutel, Paul Covington, Sagar Jain, Francois Belletti, and Ed Chi. WSDM, 2019.
 - Semi-Parametric Efficient Policy Learning with Continuous Actions
 - Victor Chernozhukov, Mert Demirer, Greg Lewis, and Vasilis Syrgkanis. NeurIPS, 2019.
 - Efficient Counterfactual Learning from Bandit Feedback
 - Yusuke Narita, Shota Yasui, and Kohei Yata. AAAI, 2019.
 - Deep Learning with Logged Bandit Feedback
 - Thorsten Joachims, Adith Swaminathan, and Maarten de Rijke. ICLR, 2018.
 - The Self-Normalized Estimator for Counterfactual Learning
 - Adith Swaminathan and Thorsten Joachims. NeurIPS, 2015.
 - Counterfactual Risk Minimization: Learning from Logged Bandit Feedback
 - Adith Swaminathan and Thorsten Joachims. ICML, 2015.

Off-Policy Evaluation and Learning: Benchmarks/Experiments

- Towards Assessing and Benchmarking Risk-Return Tradeoff of Off-Policy Evaluation
 - Haruka Kiyohara, Ren Kishimoto, Kosuke Kawakami, Ken Kobayashi, Kazuhide Nakata, and Yuta Saito. ICLR, 2024.
- SCOPE-RL: A Python Library for Offline Reinforcement Learning and Off-Policy Evaluation
 - Haruka Kiyohara, Ren Kishimoto, Kosuke Kawakami, Ken Kobayashi, Kazuhide Nakata, and Yuta Saito. arXiv, 2023.
- Offline Policy Comparison with Confidence: Benchmarks and Baselines
 - Anurag Koul, Mariano Phielipp, and Alan Fern. arXiv, 2022.
- Extending Open Bandit Pipeline to Simulate Industry Challenges
 - Bram van den Akker, Niklas Weber, Felipe Moraes, and Dmitri Goldenberg. arXiv, 2022.

-
- Open Bandit Dataset and Pipeline: Towards Realistic and Reproducible Off-Policy Evaluation [software] [public dataset]
 - Yuta Saito, Shunsuke Aihara, Megumi Matsutani, and Yusuke Narita. NeurIPS, 2021.
 - Evaluating the Robustness of Off-Policy Evaluation [software]
 - Yuta Saito, Takuma Udagawa, Haruka Kiyohara, Kazuki Mogi, Yusuke Narita, and Kei Tateno. RecSys, 2021.
 - Benchmarks for Deep Off-Policy Evaluation [code]
 - Justin Fu, Mohammad Norouzi, Ofir Nachum, George Tucker, Ziyu Wang, Alexander Novikov, Mengjiao Yang, Michael R Zhang, Yutian Chen, Aviral Kumar, Cosmin Paduraru, Sergey Levine, and Thomas Paine. ICLR, 2021.
 - Empirical Study of Off-Policy Policy Evaluation for Reinforcement Learning [code]
 - Cameron Voloshin, Hoang M. Le, Nan Jiang, and Yisong Yue, arXiv, 2019.

Off-Policy Evaluation and Learning: Applications

- HOPE: Human-Centric Off-Policy Evaluation for E-Learning and Healthcare
 - Ge Gao, Song Ju, Markel Sanz Ausin, and Min Chi. AAMAS, 2023.
- When is Off-Policy Evaluation Useful? A Data-Centric Perspective
 - Hao Sun, Alex J. Chan, Nabeel Seedat, Alihan Hüyük, and Mihaela van der Schaar. arXiv, 2023.
- Counterfactual Evaluation of Peer-Review Assignment Policies
 - Martin Saveski, Steven Jecmen, Nihar B. Shah, and Johan Ugander. arXiv, 2023.
- Balanced Off-Policy Evaluation for Personalized Pricing
 - Adam N. Elmachtoub, Vishal Gupta, and Yunfan Zhao. arXiv, 2023.
- Multi-Action Dialog Policy Learning from Logged User Feedback
 - Shuo Zhang, Junzhou Zhao, Pinghui Wang, Tianxiang Wang, Zi Liang, Jing Tao, Yi Huang, and Junlan Feng. arXiv, 2023.
- CFR-p: Counterfactual Regret Minimization with Hierarchical Policy Abstraction, and its Application to Two-player Mahjong
 - Shiheng Wang. arXiv, 2023.

-
- Reward Shaping for User Satisfaction in a REINFORCE Recommender
 - Konstantina Christakopoulou, Can Xu, Sai Zhang, Sriraj Badam, Trevor Potter, Daniel Li, Hao Wan, Xinyang Yi, Ya Le, Chris Berg, Eric Bencomo Dixon, Ed H. Chi, and Minmin Chen. arXiv, 2022.
 - Data-Driven Off-Policy Estimator Selection: An Application in User Marketing on An Online Content Delivery Service
 - Yuta Saito, Takuma Udagawa, and Kei Tateno. arXiv, 2021.
 - Towards Automatic Evaluation of Dialog Systems: A Model-Free Off-Policy Evaluation Approach
 - Haoming Jiang, Bo Dai, Mengjiao Yang, Wei Wei, and Tuo Zhao. arXiv, 2021.
 - Model Selection for Offline Reinforcement Learning: Practical Considerations for Healthcare Settings
 - Shengpu Tang and Jenna Wiens. MLHC, 2021.
 - Off-Policy Evaluation of Probabilistic Identity Data in Lookalike Modeling
 - Randell Cotta, Dan Jiang, Mingyang Hu, and Peizhou Liao. WSDM, 2019.
 - Offline Evaluation to Make Decisions About Playlist Recommendation
 - Alois Gruson, Praveen Chandar, Christophe Charbuillet, James McInerney, Samantha Hansen, Damien Tardieu, and Ben Carterette. WSDM, 2019.
 - Behaviour Policy Estimation in Off-Policy Policy Evaluation: Calibration Matters
 - Aniruddh Raghu, Omer Gottesman, Yao Liu, Matthieu Komorowski, Aldo Faisal, Finale Doshi-Velez, and Emma Brunskill. arXiv, 2018.
 - Evaluating Reinforcement Learning Algorithms in Observational Health Settings
 - Omer Gottesman, Fredrik Johansson, Joshua Meier, Jack Dent, Donghun Lee, Srivatsan Srinivasan, Linying Zhang, Yi Ding, David Wihl, Xuefeng Peng, Jiayu Yao, Isaac Lage, Christopher Mosch, Li-wei H. Lehman, Matthieu Komorowski, Matthieu Komorowski, Aldo Faisal, Leo Anthony Celi, David Sontag, and Finale Doshi-Velez. arXiv, 2018.
 - Towards a Fair Marketplace: Counterfactual Evaluation of the trade-off between Relevance, Fairness & Satisfaction in Recommendation Systems
 - Rishabh Mehrotra, James McInerney, Hugues Bouchard, Mounia Lalmas, and Fernando Diaz. CIKM, 2018.
 - Offline A/B testing for Recommender Systems

-
- Alexandre Gilotte, Clément Calauzènes, Thomas Nedelec, Alexandre Abraham, and Simon Dollé. WSDM, 2018.
 - Offline Comparative Evaluation with Incremental, Minimally-Invasive Online Feedback
 - Ben Carterette and Praveen Chandar. SIGIR, 2018.
 - Handling Confounding for Realistic Off-Policy Evaluation
 - Saurabh Sohoney, Nikita Prabhu, and Vineet Chaoji. WWW, 2018.
 - Counterfactual Reasoning and Learning Systems: The Example of Computational Advertising
 - Léon Bottou, Jonas Peters, Joaquin Quiñonero-Candela, Denis X. Charles, D. Max Chickering, Elon Portugaly, Dipankar Ray, Patrice Simard, and Ed Snelson. JMLR, 2013.

Open Source Software/Implementations

- SCOPE-RL: A Python library for offline reinforcement learning, off-policy evaluation, and selection [paper1] [paper2] [documentation]
 - Haruka Kiyohara, Ren Kishimoto, Kosuke Kawakami, Ken Kobayashi, Kazuhide Nakata, and Yuta Saito.
- Open Bandit Pipeline: a research framework for bandit algorithms and off-policy evaluation [paper] [documentation] [dataset]
 - Yuta Saito, Shunsuke Aihara, Megumi Matsutani, and Yusuke Narita.
- pylEOE: Towards An Interpretable Evaluation for Offline Evaluation [paper]
 - Yuta Saito, Takuma Udagawa, Haruka Kiyohara, Kazuki Mogi, Yusuke Narita, and Kei Tateno.
- d3rlpy: An Offline Deep Reinforcement Learning Library [paper] [website] [documentation]
 - Takuma Seno and Michita Imai.
- MINERVA: An out-of-the-box GUI tool for data-driven deep reinforcement learning [website] [documentation]
 - Takuma Seno and Michita Imai.
- Minari
 - Farama Foundation.
- CORL: Clean Offline Reinforcement Learning [paper]

-
- Denis Tarasov, Alexander Nikulin, Dmitry Akimov, Vladislav Kurenkov, and Sergey Kolesnikov.
 - COBS: Caltech OPE Benchmarking Suite [paper]
 - Cameron Voloshin, Hoang M. Le, Nan Jiang, and Yisong Yue.
 - Benchmarks for Deep Off-Policy Evaluation [paper]
 - Justin Fu, Mohammad Norouzi, Ofir Nachum, George Tucker, Ziyu Wang, Alexander Novikov, Mengjiao Yang, Michael R Zhang, Yutian Chen, Aviral Kumar, Cosmin Paduraru, Sergey Levine, and Thomas Paine.
 - DICE: The DIstribution Correction Estimation Library [paper]
 - Ofir Nachum, Yinlam Chow, Bo Dai, Lihong Li, Ruiyi Zhang, Dale Schuurmans.
 - RL Unplugged: Benchmarks for Offline Reinforcement Learning [paper] [dataset]
 - Caglar Gulcehre, Ziyu Wang, Alexander Novikov, Tom Le Paine, Sergio Gomez Colmenarejo, Konrad Zolna, Rishabh Agarwal, Josh Merel, Daniel Mankowitz, Cosmin Paduraru, Gabriel Dulac-Arnold, Jerry Li, Mohammad Norouzi, Matt Hoffman, Ofir Nachum, George Tucker, Nicolas Heess, and Nando de Freitas.
 - D4RL: Datasets for Deep Data-Driven Reinforcement Learning [paper] [website]
 - Justin Fu, Aviral Kumar, Ofir Nachum, George Tucker, and Sergey Levine.
 - V-D4RL: Challenges and Opportunities in Offline Reinforcement Learning from Visual Observations [paper]
 - Cong Lu, Philip J. Ball, Tim G. J. Rudner, Jack Parker-Holder, Michael A. Osborne, and Yee Whye Teh.
 - Benchmarking Offline Reinforcement Learning on Real-Robot Hardware [paper]
 - Nico Gürtler, Sebastian Blaes, Pavel Kolev, Felix Widmaier, Manuel Wuthrich, Stefan Bauer, Bernhard Schölkopf, and Georg Martius. ICLR, 2023.
 - RLDS: Reinforcement Learning Datasets [paper]
 - Sabela Ramos, Sertan Girgin, Léonard Hussenot, Damien Vincent, Hanna Yakubovich, Daniel Toyama, Anita Gergely, Piotr Stanczyk, Raphael Marinier, Jeremiah Harmsen, Olivier Pietquin, and Nikola Momchev.
 - OEF: Offline Equilibrium Finding [paper]
 - Shuxin Li, Xinrun Wang, Jakub Cerny, Youzhi Zhang, Hau Chan, and Bo An.
-

-
- ExORL: Exploratory Data for Offline Reinforcement Learning [paper]
 - Denis Yarats, David Brandfonbrener, Hao Liu, Michael Laskin, Pieter Abbeel, Alessandro Lazaric, and Lerrel Pinto.
 - RL4RS: A Real-World Benchmark for Reinforcement Learning based Recommender System [paper] dataset]
 - Kai Wang, Zhene Zou, Yue Shang, Qilin Deng, Minghao Zhao, Yile Liang, Runze Wu, Jianrong Tao, Xudong Shen, Tangjie Lyu, and Changjie Fan.
 - NeoRL: Near Real-World Benchmarks for Offline Reinforcement Learning [paper] [website]
 - Rongjun Qin, Songyi Gao, Xingyuan Zhang, Zhen Xu, Shengkai Huang, Zewen Li, Weinan Zhang, and Yang Yu.
 - The Industrial Benchmark Offline RL Datasets [paper]
 - Phillip Swazinna, Steffen Udluft, and Thomas Runkler.
 - ARLO: A Framework for Automated Reinforcement Learning [paper]
 - Marco Mussi, Davide Lombarda, Alberto Maria Metelli, Francesco Trovò, and Marcello Restelli.
 - RecoGym: A Reinforcement Learning Environment for the problem of Product Recommendation in Online Advertising [paper]
 - David Rohde, Stephen Bonner, Travis Dunlop, Flavian Vasile, and Alexandros Karatzoglou.
 - MARS-Gym: A Gym framework to model, train, and evaluate Recommender Systems for Marketplaces [paper] [documentation]
 - Marlesson R. O. Santana, Luckeciano C. Melo, Fernando H. F. Camargo, Bruno Brandão, Anderson Soares, Renan M. Oliveira, and Sandor Caetano.
 - A Reinforcement Learning-based Volt-VAR Control Dataset [paper]
 - Yuanqi Gao and Nanpeng Yu.

Blog/Podcast

Blog

- Counterfactual Evaluation for Recommendation Systems
 - Eugene Yan. 2022.

-
- Offline Reinforcement Learning: How Conservative Algorithms Can Enable New Applications
 - Aviral Kumar and Avi Singh. BAIR Blog, 2020.
 - AWAC: Accelerating Online Reinforcement Learning with Offline Datasets
 - Ashvin Nair and Abhishek Gupta. BAIR Blog, 2020.
 - D4RL: Building Better Benchmarks for Offline Reinforcement Learning
 - Justin Fu. BAIR Blog, 2020.
 - Does On-Policy Data Collection Fix Errors in Off-Policy Reinforcement Learning?
 - Aviral Kumar and Abhishek Gupta. BAIR Blog, 2020.
 - Tackling Open Challenges in Offline Reinforcement Learning
 - George Tucker and Sergey Levine. Google AI Blog, 2020.
 - An Optimistic Perspective on Offline Reinforcement Learning
 - Rishabh Agarwal and Mohammad Norouzi. Google AI Blog, 2020.
 - Decisions from Data: How Offline Reinforcement Learning Will Change How We Use Machine Learning
 - Sergey Levine. Medium, 2020.
 - Introducing completely free datasets for data-driven deep reinforcement learning
 - Takuma Seno. towards data science, 2020.
 - Offline (Batch) Reinforcement Learning: A Review of Literature and Applications
 - Daniel Seita. danieltakeshi.github.io, 2020.
 - Data-Driven Deep Reinforcement Learning
 - Aviral Kumar. BAIR Blog, 2019.

Podcast

- AI Trends 2023: Reinforcement Learning – RLHF, Robotic Pre-Training, and Offline RL with Sergey Levine
 - Sergey Levine. TWIML, 2023.
- Bandits and Simulators for Recommenders with Olivier Jeunen

-
- Olivier Jeunen. Recsperts, 2022.
 - Sergey Levine on Robot Learning & Offline RL
 - Sergey Levine. The Gradient, 2021.
 - Off-Line, Off-Policy RL for Real-World Decision Making at Facebook
 - Jason Gauci. TWIML, 2021.
 - Xianyuan Zhan | TalkRL: The Reinforcement Learning Podcast
 - Xianyuan Zhan. TWIML, 2021.
 - MOREL: Model-Based Offline Reinforcement Learning with Aravind Rajeswaran
 - Aravind Rajeswaran. TWIML, 2020.
 - Trends in Reinforcement Learning with Chelsea Finn
 - Chelsea Finn. TWIML, 2020.
 - Nan Jiang | TalkRL: The Reinforcement Learning Podcast
 - Nan Jiang. TalkRL, 2020.
 - Scott Fujimoto | TalkRL: The Reinforcement Learning Podcast
 - Scott Fujimoto. TalkRL, 2019.

Related Workshops

- CONSEQUENCES (RecSys 2023)
- Offline Reinforcement Learning (NeurIPS 2022)
- Reinforcement Learning for Real Life (NeurIPS 2022)
- CONSEQUENCES + REVEAL (RecSys 2022)
- Offline Reinforcement Learning (NeurIPS 2021)
- Reinforcement Learning for Real Life (ICML 2021)
- Reinforcement Learning Day 2021
- Offline Reinforcement Learning (NeurIPS 2020)
- Reinforcement Learning from Batch Data and Simulation
- Reinforcement Learning for Real Life (RL4RealLife 2020)
- Safety and Robustness in Decision Making (NeurIPS 2019)
- Reinforcement Learning for Real Life (ICML 2019)
- Real-world Sequential Decision Making (ICML 2019)

Tutorials/Talks/Lectures

- Reinforcement Learning with Large Datasets: Robotics, Image Generation, and LLMs
 - Sergey Levine. 2023.
- Counterfactual Evaluation and Learning for Interactive Systems
 - Yuta Saito and Thorsten Joachims. KDD2022.
- Representation Learning for Online and Offline RL in Low-rank MDPs
 - Masatoshi Uehara. RL Theory Seminar2022.
- Offline Reinforcement Learning: Fundamental Barriers for Value Function Approximation
 - Yunzong Xu. RL Theory Seminar2022.
- Safe Policy Learning through Extrapolation: Application to Pre-trial Risk Assessment
 - Kosuke Imai. Online Causal Inference Seminar2022.
- Deep Reinforcement Learning with Real-World Data
 - Sergey Levine. 2022.
- Planning with Reinforcement Learning
 - Sergey Levine. 2022.
- Imitation learning vs. offline reinforcement learning
 - Sergey Levine. 2022.
- Tutorial on the Foundations of Offline Reinforcement Learning
 - Romain Laroche and David Brandfonbrener. 2022.
- Counterfactual Learning and Evaluation for Recommender Systems: Foundations, Implementations, and Recent Advances [website]
 - Yuta Saito and Thorstem Joachims. RecSys2021.
- Offline Reinforcement Learning
 - Sergey Levine. BayLearn2021.
- Offline Reinforcement Learning
 - Guy Tennenholtz. CHIL2021.

-
- Fast Rates for the Regret of Offline Reinforcement Learning
 - Yichun Hu. RL Theory Seminar2021.
 - Bellman-consistent Pessimism for Offline Reinforcement Learning
 - Tengyan Xie. RL Theory Seminar2021.
 - Pessimistic Model-based Offline Reinforcement Learning under Partial Coverage
 - Masatoshi Uehara. RL Theory Seminar2021.
 - Bridging Offline Reinforcement Learning and Imitation Learning: A Tale of Pessimism
 - Paria Rashidinejad. RL Theory Seminar2021.
 - Infinite-Horizon Offline Reinforcement Learning with Linear Function Approximation: Curse of Dimensionality and Algorithm
 - Lin Chen. RL Theory Seminar2021.
 - Is Pessimism Provably Efficient for Offline RL?
 - Ying Jin. RL Theory Seminar2021.
 - Adaptive Estimator Selection for Off-Policy Evaluation
 - Yi Su. RL Theory Seminar2021.
 - What are the Statistical Limits of Offline RL with Linear Function Approximation?
 - Ruosong Wang. RL Theory Seminar2021.
 - Exponential Lower Bounds for Batch Reinforcement Learning: Batch RL can be Exponentially Harder than Online RL
 - Andrea Zanette. RL Theory Seminar2021.
 - A Gentle Introduction to Offline Reinforcement Learning
 - Sergey Levine. 2021.
 - Principles for Tackling Distribution Shift: Pessimism, Adaptation, and Anticipation
 - Chelsea Finn. 2020-2021 Machine Learning Advances and Applications Seminar.
 - Offline Reinforcement Learning: Incorporating Knowledge from Data into RL
 - Sergey Levine. IJCAI-PRICAI2020 Knowledge Based Reinforcement Learning Workshop.
 - Offline RL

-
- Nando de Freitas. NeurIPS2020 OfflineRL Workshop.
 - Learning a Multi-Agent Simulator from Offline Demonstrations
 - Brandyn White. NeurIPS2020 OfflineRL Workshop.
 - Towards Reliable Validation and Evaluation for Offline RL
 - Nan Jiang. NeurIPS2020 OfflineRL Workshop.
 - Batch RL Models Built for Validation
 - Finale Doshi-Velez. NeurIPS2020 OfflineRL Workshop.
 - Offline Reinforcement Learning: From Algorithms to Practical Challenges
 - Aviral Kumar and Sergey Levine. NeurIPS2020.
 - Data Scalability for Robot Learning
 - Chelsea Finn. RL Seminar2020.
 - Statistically Efficient Offline Reinforcement Learning
 - Nathan Kallus. ARL Seminar2020.
 - Near Optimal Provable Uniform Convergence in Off-Policy Evaluation for Reinforcement Learning
 - Yu-Xiang Wang. RL Theory Seminar2020.
 - Minimax-Optimal Off-Policy Evaluation with Linear Function Approximation
 - Mengdi Wang. RL Theory Seminar2020.
 - Beyond the Training Distribution: Embodiment, Adaptation, and Symmetry
 - Chelsea Finn. EL Seminar2020.
 - Combining Statistical methods with Human Input for Evaluation and Optimization in Batch Settings
 - Finale Doshi-Velez. NeurIPS2019 Workshop on Safety and Robustness in Decision Making.
 - Efficiently Breaking the Curse of Horizon with Double Reinforcement Learning
 - Nathan Kallus. NeurIPS2019 Workshop on Safety and Robustness in Decision Making.
 - Scaling Probabilistically Safe Learning to Robotics
 - Scott Niekum. NeurIPS2019 Workshop on Safety and Robustness in Decision Making.
 - Deep Reinforcement Learning in the Real World
 - Sergey Levine. Workshop on New Directions in Reinforcement Learning and Control2019.
-