1. Record Nr. UNISA996464515003316

| | |
|---|---|
| Titolo | Machine learning and knowledge discovery in databases . Part I : Research track : European Conference, ECML PKDD 2021, Bilbao, Spain, September 13-17, 2021, Proceedings / / Nuria Oliver [and four others], (editors) |
| Pubbl/distr/stampa | Cham, Switzerland : , : Springer, , [2021]<br>©2021 |
| ISBN | 3-030-86486-3 |
| Descrizione fisica | 1 online resource (838 pages) |
| Collana | Lecture notes in computer science ; ; 12975 |
| Disciplina | 006.31 |
| Soggetti | Machine learning |
| Lingua di pubblicazione | Inglese |
| Formato | Materiale a stampa |
| Livello bibliografico | Monografia |
| Nota di bibliografia | Includes bibliographical references and index. |
| Nota di contenuto | Intro -- Preface -- Organization -- Invited Talks Abstracts -- WuDao: Pretrain the World -- The Value of Data for Personalization -- AI Fairness in Practice -- Safety and Robustness for Deep Learning with Provable Guarantees -- Contents - Part I -- Online Learning -- Routine Bandits: Minimizing Regret on Recurring Problems -- 1 Introduction -- 2 The Routine Bandit Setting -- 3 The KLUCB-RB Strategy -- 4 Sketch of Proof -- 5 Numerical Experiments -- 5.1 More Arms Than Bandits: A Beneficial Case -- 5.2 Increasing the Number of Bandit Instances -- 5.3 Critical Settings -- 6 Conclusion -- References -- Conservative Online Convex Optimization -- 1 Introduction -- 2 Background -- 3 Problem Formulation -- 4 The Conservative Projection Algorithm -- 4.1 The Conservative Ball -- 4.2 Description of the CP Algorithm -- 4.3 Analysis of the CP Algorithm -- 5 Experiments -- 5.1 Synthetic Regression Dataset -- 5.2 Online Classification: The IMDB Dataset -- 5.3 Online Classification: The SpamBase Dataset -- 6 Conclusions -- References -- Knowledge Infused Policy Gradients with Upper Confidence Bound for Relational Bandits -- 1 Introduction -- 2 Problem Setting -- 3 Knowledge Infused Policy Gradients -- 4 Formulation of Knowledge Infusion -- 5 Regret Bound for KIPG -- 6 KIPG-Upper Confidence Bound -- 7 Experiments -- 7.1 Simulated Domains -- 7.2 Real-World Datasets -- 8 Conclusion and Future Work -- References -- |

Exploiting History Data for Nonstationary Multi-armed Bandit -- 1 Introduction -- 2 Related Works -- 3 Problem Formulation -- 4 The BR-MAB Algorithm -- 4.1 Break-Point Prediction Procedure -- 4.2 Recurrent Concepts Equivalence Test -- 4.3 Regret Analysis for Generic CD-MABs -- 4.4 Regret Analysis for the Break-Point Prediction Procedure -- 5 Experiments -- 5.1 Toy Example -- 5.2 Synthetic Setting -- 5.3 Yahoo! Setting -- 6 Conclusion and Future Works. References -- High-Probability Kernel Alignment Regret Bounds for Online Kernel Selection -- 1 Introduction -- 1.1 Related Work -- 2 Problem Setting -- 3 A Nearly Optimal High-Probability Regret Bound -- 3.1 Warm-Up -- 3.2 A More Efficient Algorithm -- 3.3 Regret Bound -- 3.4 Time Complexity Analysis -- 4 Regret-Performance Trade-Off -- 4.1 Regret Bound -- 4.2 Budgeted EA2OKS -- 5 Experiments -- 5.1 Experimental Setting -- 5.2 Experimental Results -- 6 Conclusion -- References -- Reinforcement Learning -- Periodic Intra-ensemble Knowledge Distillation for Reinforcement Learning -- 1 Introduction -- 2 Related Work -- 3 Background -- 4 Method -- 4.1 Overview -- 4.2 Ensemble Initialization -- 4.3 Joint Training -- 4.4 Intra-ensemble Knowledge Distillation -- 5 Experiments -- 5.1 Experimental Setup -- 5.2 Effectiveness of PIEKD -- 5.3 Effectiveness of Knowledge Distillation for Knowledge Sharing -- 5.4 Effectiveness of Selecting the Best-Performing Agent as the Teacher -- 5.5 Ablation Study on Ensemble Size -- 5.6 Ablation Study on Distillation Interval -- 6 Conclusion -- References -- Learning to Build High-Fidelity and Robust Environment Models -- 1 Introduction -- 2 Related Work -- 2.1 Simulator Building -- 2.2 Model-Based Reinforcement Learning -- 2.3 Offline Policy Evaluation -- 2.4 Robust Reinforcement Learning -- 3 Preliminaries -- 3.1 Markov Decision Process -- 3.2 Dual Markov Decision Process -- 3.3 Imitation Learning -- 4 Robust Learning to Simulate -- 4.1 Problem Definition -- 4.2 Single Behavior Policy Setting -- 4.3 Robust Policy Setting -- 5 Experiments -- 5.1 Experimental Protocol -- 5.2 Studied Environments and Baselines -- 5.3 Performance on Policy Value Difference Evaluation -- 5.4 Performance on Policy Ranking -- 5.5 Performance on Policy Improvement -- 5.6 Analysis on Hyperparameter -- 6 Conclusion -- References.
Ensemble and Auxiliary Tasks for Data-Efficient Deep Reinforcement Learning -- 1 Introduction -- 2 Related Works -- 3 Background -- 3.1 Markov Decision Process and RL -- 3.2 Rainbow Agent -- 4 Rainbow Ensemble -- 5 Auxiliary Tasks for Ensemble RL -- 5.1 Network Architecture -- 5.2 Model Learning as Auxiliary Tasks -- 5.3 Object and Event Based Auxiliary Tasks -- 6 Theoretical Analysis -- 7 Experiments -- 7.1 Comparison to Prior Works -- 7.2 Bias-Variance-Covariance Measurements -- 7.3 On Independent Training of Ensemble -- 7.4 The Importance of Auxiliary Tasks -- 7.5 On Distributing the Auxiliary Tasks -- 8 Conclusions -- References -- Multi-agent Imitation Learning with Copulas -- 1 Introduction -- 2 Preliminaries -- 3 Modeling Multi-agent Interaction with Copulas -- 3.1 Copulas -- 3.2 Multi-agent Imitation Learning with Copulas -- 4 Related Work -- 5 Experiments -- 5.1 Experimental Setup -- 5.2 Results -- 5.3 Generalization of Copula -- 5.4 Copula Visualization -- 5.5 Trajectory Generation -- 6 Conclusion and Future Work -- A  Dataset Details -- B Implementation Details -- References -- CMIX: Deep Multi-agent Reinforcement Learning with Peak and Average Constraints -- 1 Introduction -- 2 Background -- 2.1 QMIX -- 2.2 Constrained Reinforcement Learning -- 3 Problem Formulation -- 4 CMIX -- 4.1 Multi-objective Constrained Problem -- 4.2 CMIX Architecture -- 4.3 Gap Loss Function -- 4.4 CMIX Algorithm -- 5 Experiments -- 5.1 Blocker Game with Travel Cost -- 5.2 Vehicular Network Routing

Optimization -- 5.3 Gap Loss Coefficient -- 6 Related Work -- 7 Conclusion -- References -- Model-Based Offline Policy Optimization with Distribution Correcting Regularization -- 1 Introduction -- 2 Preliminary -- 2.1 Markov Decision Processes -- 2.2 Offline RL -- 2.3 Model-Based RL -- 3 A Lower Bound of the True Expected Return -- 4 Method -- 4.1 Overall Framework.
4.2 Ratio Estimation via DICE -- 5 Experiment -- 5.1 Comparative Evaluation -- 5.2 Empirical Analysis -- 6 Related Work -- 6.1 Model-Free Offline RL -- 6.2 Model-Based Offline RL -- 7 Conclusion -- References -- Disagreement Options: Task Adaptation Through Temporally Extended Actions -- 1 Introduction -- 2 Preliminaries -- 3 Disagreement Options -- 3.1 Task Similarity: How to Select Relevant Priors? -- 3.2 Task Adaptation: How Should We Use the Prior Knowledge? -- 3.3 Prior Policy Acquisition -- 4 Experiments -- 4.1 3D MiniWorld -- 4.2 Photorealistic Simulator -- 5 Towards Real-World Task Adaptation -- 6 Related Work -- 7 Discussion -- 8 Conclusion -- References -- Deep Adaptive Multi-intention Inverse Reinforcement Learning -- 1 Introduction -- 2 Related Works -- 3 Problem Definition -- 4 Approach -- 4.1 First Solution with Stochastic Expectation Maximization -- 4.2 Second Solution with Monte Carlo Expectation Maximization -- 5 Experimental Results -- 5.1 Benchmarks -- 5.2 Models -- 5.3 Metric -- 5.4 Implementations Details -- 5.5 Results -- 6 Conclusions -- References -- Unsupervised Task Clustering for Multi-task Reinforcement Learning -- 1 Introduction -- 2 Related Work -- 3 Background and Notation -- 4 Clustered Multi-task Learning -- 4.1 Convergence Analysis -- 5 Experiments -- 5.1 Pendulum -- 5.2 Bipedal Walker -- 5.3 Atari -- 5.4 Ablations -- 6 Conclusion -- References -- Deep Model Compression via Two-Stage Deep Reinforcement Learning -- 1 Introduction -- 2 A Deep Reinforcement Learning Compression Framework -- 2.1 State -- 2.2 Action -- 2.3 Reward -- 2.4 The Proposed DRL Compression Structure -- 3 Pruning -- 3.1 Pruning from C Dimension: Channel Pruning -- 3.2 Pruning from H and W Dimensions: Variational Pruning -- 4 Quantization -- 5 Experiments -- 5.1 Settings -- 5.2 MNIST and CIFAR-10 -- 5.3 ImageNet.
5.4 Variational Pruning via Information Dropout -- 5.5 Single Layer Acceleration Performance -- 5.6 Time Complexity -- 6 Conclusion -- References -- Dropout's Dream Land: Generalization from Learned Simulators to Reality -- 1 Introduction -- 2 Related Works -- 2.1 Dropout -- 2.2 Domain Randomization -- 2.3 World Models -- 3 Dropout's Dream Land -- 3.1 Learning the Dream Environment -- 3.2 Interacting with Dropout's Dream Land -- 3.3 Training the Controller -- 4 Experiments -- 4.1 Comparison with Baselines -- 4.2 Inference Dropout and Dream2Real Generalization -- 4.3 When Should Dropout Masks Be Randomized During Controller Training? -- 4.4 Comparison to Standard Regularization Methods -- 4.5 Comparison to Explicit Ensemble Methods -- 5 Conclusion -- References -- Goal Modelling for Deep Reinforcement Learning Agents -- 1 Introduction -- 2 Background -- 3 Deep Reinforcement Learning with Goal Net -- 4 Experiments -- 4.1 Two Keys -- 4.2 3D Four Rooms with Subgoals -- 4.3 Kitchen Navigation and Interaction -- 5 Related Work -- 6 Discussion and Conclusion -- References -- Time Series, Streams, and Sequence Models -- Deviation-Based Marked Temporal Point Process for Marker Prediction -- 1 Introduction -- 2 Related Work -- 3 Proposed Algorithm -- 3.1 Problem Definition -- 3.2 Preliminaries -- 3.3 Proposed Deviation-Based Marked Temporal Point Process -- 3.4 Implementation Details -- 4 Experiments and Protocols -- 5 Results and Analysis -- 6 Conclusion and Discussion -- References -- Deep