

35th AAAI Conference on Artificial Intelligence (AAAI-21)

Volume 11: AAAI Technical Tracks

- Machine Learning IV

Online
2 – 9 February 2021

Part 1 of 2

ISBN: 978-1-7138-3597-4

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2021) by Association for the Advancement of Artificial Intelligence
All rights reserved.

Printed with permission by Curran Associates, Inc. (2021)

For permission requests, please contact Association for the Advancement of Artificial Intelligence
at the address below.

Association for the Advancement of Artificial Intelligence
2275 East Bayshore Road
Suite 160
Palo Alto, California 94303
USA

Phone: 1-650-328-3123
Fax: 1-650-321-4457

<https://aaai.org/Press/press.php>

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2633
Email: curran@proceedings.com
Web: www.proceedings.com

TABLE OF CONTENTS

PART 1

AAAI TECHNICAL TRACK ON MACHINE LEARNING IV

AUTODROPOUT: LEARNING DROPOUT PATTERNS TO REGULARIZE DEEP NETWORKS	9351
<i>Hieu Pham, Quoc Le</i>	
FAST MULTI-VIEW DISCRETE CLUSTERING WITH ANCHOR GRAPHS	9360
<i>Qianyao Qiang, Bin Zhang, Fei Wang, Feiping Nie</i>	
RELATION-AWARE GRAPH ATTENTION MODEL WITH ADAPTIVE SELF-ADVERSARIAL TRAINING	9368
<i>Xiao Qin, Nasrullah Sheikh, Berthold Reinwald, Lingfei Wu</i>	
UNCERTAINTY-AWARE POLICY OPTIMIZATION: A ROBUST, ADAPTIVE TRUST REGION APPROACH	9377
<i>James Queeney, Ioannis Ch. Paschalidis, Christos G. Cassandras</i>	
CLASSIFYING SEQUENCES OF EXTREME LENGTH WITH CONSTANT MEMORY APPLIED TO MALWARE DETECTION	9386
<i>Edward Raff, William Fleshman, Richard Zak, Hyrum S. Anderson, Bobby Filar, Mark McLean</i>	
ONLINE DR-SUBMODULAR MAXIMIZATION: MINIMIZING REGRET AND CONSTRAINT VIOLATION	9395
<i>Prasanna Raut, Omid Sadeghi, Maryam Fazel</i>	
IMPROVING GENERATIVE MOMENT MATCHING NETWORKS WITH DISTRIBUTION PARTITION	9403
<i>Yong Ren, Yucen Luo, Jun Zhu</i>	
MULTIPLE KERNEL CLUSTERING WITH KERNEL K-MEANS COUPLED GRAPH TENSOR LEARNING	9411
<i>Zhenwen Ren, Quansen Sun, Dong Wei</i>	
ROBUST FAIRNESS UNDER COVARIATE SHIFT	9419
<i>Ashkan Rezaei, Anqi Liu, Omid Memarrast, Brian D. Ziebart</i>	
SHUFFLING RECURRENT NEURAL NETWORKS	9428
<i>Michael Rotman, Lior Wolf</i>	
WHY ADVERSARIAL INTERACTION CREATES NON-HOMOGENEOUS PATTERNS: A PSEUDO-REACTION-DIFFUSION MODEL FOR TURING INSTABILITY	9436
<i>Litu Rout</i>	
ADVERSARIAL PERMUTATION GUIDED NODE REPRESENTATIONS FOR LINK PREDICTION	9445
<i>Indradyumna Roy, Abir De, Soumen Chakrabarti</i>	

VISUAL TRANSFER FOR REINFORCEMENT LEARNING VIA WASSERSTEIN DOMAIN CONFUSION	9454
<i>Josh Roy, George D. Konidakis</i>	
ANYTIME INFERENCE WITH DISTILLED HIERARCHICAL NEURAL ENSEMBLES	9463
<i>Adria Ruiz, Jakob Verbeek</i>	
INVERSE REINFORCEMENT LEARNING WITH EXPLICIT POLICY ESTIMATES	9472
<i>Navyata Sanghvi, Shinnosuke Usami, Mohit Sharma, Joachim Groeger, Kris Kitani</i>	
A DEEPER LOOK AT THE HESSIAN EIGENSPECTRUM OF DEEP NEURAL NETWORKS AND ITS APPLICATIONS TO REGULARIZATION	9481
<i>Adepu Ravi Sankar, Yash Khasbage, Rahul Vigneswaran, Vineeth N Balasubramanian</i>	
ADVANTAGENAS: EFFICIENT NEURAL ARCHITECTURE SEARCH WITH CREDIT ASSIGNMENT	9489
<i>Rei Sato, Jun Sakuma, Youhei Akimoto</i>	
ACTIVE FEATURE SELECTION FOR THE MUTUAL INFORMATION CRITERION	9497
<i>Shachar Schnapp, Sivan Sabato</i>	
LEARNING PRECISE TEMPORAL POINT EVENT DETECTION WITH MISALIGNED LABELS	9505
<i>Julien Schroeter, Kirill Sidorov, David Marshall</i>	
MULTI-TYPE DISENTANGLEMENT WITHOUT ADVERSARIAL TRAINING	9515
<i>Lei Sha, Thomas Lukasiewicz</i>	
UNCERTAINTY-MATCHING GRAPH NEURAL NETWORKS TO DEFEND AGAINST POISONING ATTACKS	9524
<i>Uday Shankar Shanthamallu, Jayaraman J. Thiagarajan, Andreas Spanias</i>	
RIGHT FOR BETTER REASONS: TRAINING DIFFERENTIABLE MODELS BY CONSTRAINING THEIR INFLUENCE FUNCTIONS	9533
<i>Xiaoting Shao, Arseny Skryagin, Wolfgang Stammer, Patrick Schramowski, Kristian Kersting</i>	
META-LEARNING EFFECTIVE EXPLORATION STRATEGIES FOR CONTEXTUAL BANDITS	9541
<i>Amr Sharaf, Hal Daumé III</i>	
MEMBERSHIP PRIVACY FOR MACHINE LEARNING MODELS THROUGH KNOWLEDGE TRANSFER	9549
<i>Virat Shejwalkar, Amir Houmansadr</i>	
THEORETICALLY PRINCIPLED DEEP RL ACCELERATION VIA NEAREST NEIGHBOR FUNCTION APPROXIMATION	9558
<i>Junhong Shen, Lin F. Yang</i>	
TIME SERIES ANOMALY DETECTION WITH MULTIREOLUTION ENSEMBLE DECODING	9567
<i>Lifeng Shen, Zhongzhong Yu, Qianli Ma, James T. Kwok</i>	
STL-SGD: SPEEDING UP LOCAL SGD WITH STAGewise COMMUNICATION PERIOD	9576
<i>Shuheng Shen, Yifei Cheng, Jingchang Liu, Linli Xu</i>	

PDO-ES2CNNS: PARTIAL DIFFERENTIAL OPERATOR BASED EQUIVARIANT SPHERICAL CNNs	9585
<i>Zhengyang Shen, Tiancheng Shen, Zhouchen Lin, Jinwen Ma</i>	
PARTIAL IS BETTER THAN ALL: REVISITING FINE-TUNING STRATEGY FOR FEW-SHOT LEARNING	9594
<i>Zhiqiang Shen, Zechun Liu, Jie Qin, Marios Savvides, Kwang-Ting Cheng</i>	
FEDERATED MULTI-ARMED BANDITS	9603
<i>Chengshuai Shi, Cong Shen</i>	
RAVEN'S PROGRESSIVE MATRICES COMPLETION WITH LATENT GAUSSIAN PROCESS PRIORS	9612
<i>Fan Shi, Bin Li, Xiangyang Xue</i>	
IMPROVED PENALTY METHOD VIA DOUBLY STOCHASTIC GRADIENTS FOR BILEVEL HYPERPARAMETER OPTIMIZATION	9621
<i>Wanli Shi, Bin Gu</i>	
ONLINE CLASS-INCREMENTAL CONTINUAL LEARNING WITH ADVERSARIAL SHAPLEY VALUE	9630
<i>Dongsub Shim, Zheda Mai, Jihwan Jeong, Scott Sanner, Hyunwoo Kim, Jongseong Jang</i>	
SCALABLE AFFINITY PROPAGATION FOR MASSIVE DATASETS	9639
<i>Hiroaki Shiokawa</i>	
INTERPRETABLE SEQUENCE CLASSIFICATION VIA DISCRETE OPTIMIZATION	9647
<i>Maayan Shvo, Andrew C. Li, Rodrigo Toro Icarte, Sheila A. McIlraith</i>	
TOWARDS DOMAIN INVARIANT SINGLE IMAGE DEHAZING	9657
<i>Pranjay Shyam, Kuk-Jin Yoon, Kyung-Soo Kim</i>	
DIBS: DIVERSITY INDUCING INFORMATION BOTTLENECK IN MODEL ENSEMBLES	9666
<i>Samarth Sinha, Homanga Bharadhwaj, Anirudh Goyal, Hugo Larochelle, Animesh Garg, Florian Shkurti</i>	
DIFFERENTIAL SPECTRAL NORMALIZATION (DSN) FOR PDE DISCOVERY	9675
<i>Chi Chiu So, Tsz On Li, Chufang Wu, Siu Pang Yung</i>	
UNIPOINT: UNIVERSALLY APPROXIMATING POINT PROCESSES INTENSITIES	9685
<i>Alexander Soen, Alexander Mathews, Daniel Grixti-Cheng, Lexing Xie</i>	
SOLVING COMMON-PAYOFF GAMES WITH APPROXIMATE POLICY ITERATION	9695
<i>Samuel Sokota, Edward Lockhart, Finbarr Timbers, Elnaz Davoodi, Ryan D'Orazio, Neil Burch, Martin Schmid, Michael Bowling, Marc Lanctot</i>	
IMPROVING GRADIENT FLOW WITH UNROLLED HIGHWAY EXPECTATION MAXIMIZATION	9704
<i>Chonghyuk Song, Eunseok Kim, Inwook Shim</i>	
IMPLICIT KERNEL ATTENTION	9713
<i>Kyungwoo Song, Yohan Jung, Dongjun Kim, Il-Chul Moon</i>	
ERROR-CORRECTING OUTPUT CODES WITH ENSEMBLE DIVERSITY FOR ROBUST LEARNING IN NEURAL NETWORKS	9722
<i>Yang Song, Qiyu Kang, Wee Peng Tay</i>	

HIERARCHICAL RELATIONAL INFERENCE	9730
<i>Aleksandar Stanic, Sjoerd Van Steenkiste, Jürgen Schmidhuber</i>	
'LESS THAN ONE'-SHOT LEARNING: LEARNING N CLASSES FROM $M < N$ SAMPLES	9739
<i>Iliia Sucholutsky, Matthias Schonlau</i>	
HIABP: HIERARCHICAL INITIALIZED ABP FOR UNSUPERVISED REPRESENTATION LEARNING	9747
<i>Jiankai Sun, Rui Liu, Bolei Zhou</i>	
STABILITY AND GENERALIZATION OF DECENTRALIZED STOCHASTIC GRADIENT DESCENT	9756
<i>Tao Sun, Dongsheng Li, Bao Wang</i>	
TEMPLE: LEARNING TEMPLATE OF TRANSITIONS FOR SAMPLE EFFICIENT MULTI- TASK RL	9765
<i>Yanchao Sun, Xiangyu Yin, Furong Huang</i>	
PAC LEARNING OF CAUSAL TREES WITH LATENT VARIABLES	9774
<i>Prasad Tadepalli, Stuart J. Russell</i>	
LEARNING DYNAMICS MODELS WITH STABLE INVARIANT SETS	9782
<i>Naoya Takeishi, Yoshinobu Kawahara</i>	
NEAR-OPTIMAL REGRET BOUNDS FOR CONTEXTUAL COMBINATORIAL SEMI- BANDITS WITH LINEAR PAYOFF FUNCTIONS	9791
<i>Kei Takemura, Shinji Ito, Daisuke Hatano, Hanna Sumita, Takuro Fukunaga, Naonori Kakimura, Ken-Ichi Kawarabayashi</i>	

PART 2

EXPLICITLY MODELED ATTENTION MAPS FOR IMAGE CLASSIFICATION	9799
<i>Andong Tan, Duc Tam Nguyen, Maximilian Dax, Matthias Nießner, Thomas Brox</i>	
PROXY GRAPH MATCHING WITH PROXIMAL MATCHING NETWORKS	9808
<i>Hao-Ru Tan, Chuang Wang, Si-Tong Wu, Tie-Qiang Wang, Xu-Yao Zhang, Cheng-Lin Liu</i>	
STRATEGY AND BENCHMARK FOR CONVERTING DEEP Q-NETWORKS TO EVENT- DRIVEN SPIKING NEURAL NETWORKS	9816
<i>Weihao Tan, Devdhar Patel, Robert Kozma</i>	
EMPOWERING ADAPTIVE EARLY-EXIT INFERENCE WITH LATENCY AWARENESS	9825
<i>Xinrui Tan, Hongjia Li, Liming Wang, Xueqing Huang, Zhen Xu</i>	
FORESEE THEN EVALUATE: DECOMPOSING VALUE ESTIMATION WITH LATENT FUTURE PREDICTION	9834
<i>Hongyao Tang, Zhaopeng Meng, Guangyong Chen, Pengfei Chen, Chen Chen, Yaodong Yang, Luo Zhang, Wulong Liu, Jianye Hao</i>	
GRADIENT DESCENT AVERAGING AND PRIMAL-DUAL AVERAGING FOR STRONGLY CONVEX OPTIMIZATION	9843
<i>Wei Tao, Wei Li, Zhisong Pan, Qing Tao</i>	
EVOLUTIONARY APPROACH FOR AUTOAUGMENT USING THE THERMODYNAMICAL GENETIC ALGORITHM	9851
<i>Akira Terauchi, Naoki Mori</i>	

SEMI-SUPERVISED KNOWLEDGE AMALGAMATION FOR SEQUENCE CLASSIFICATION	9859
<i>Jidapa Thadajarassiri, Thomas Hartvigsen, Xiangnan Kong, Elke A Rundensteiner</i>	
ONLINE NON-MONOTONE DR-SUBMODULAR MAXIMIZATION	9868
<i>Nguyen Kim Thang, Abhinav Srivastav</i>	
DETECTING ADVERSARIAL EXAMPLES FROM SENSITIVITY INCONSISTENCY OF SPATIAL-TRANSFORM DOMAIN	9877
<i>Jinyu Tian, Jiantao Zhou, Yuanman Li, Jia Duan</i>	
TOWARDS TRUSTWORTHY PREDICTIONS FROM DEEP NEURAL NETWORKS WITH FAST ADVERSARIAL CALIBRATION	9886
<i>Christian Tomani, Florian Buettner</i>	
META LEARNING FOR CAUSAL DIRECTION	9897
<i>Jean-François Ton, Dino Sejdinovic, Kenji Fukumizu</i>	
LEARNING COMPOSITIONAL SPARSE GAUSSIAN PROCESSES WITH A SHRINKAGE PRIOR	9906
<i>Anh Tong, Toan M Tran, Hung Bui, Jaesik Choi</i>	
CHARACTERIZING DEEP GAUSSIAN PROCESSES VIA NONLINEAR RECURRENCE SYSTEMS	9915
<i>Anh Tong, Jaesik Choi</i>	
ITERATIVE BOUNDING MDPS: LEARNING INTERPRETABLE POLICIES VIA NON-INTERPRETABLE METHODS	9923
<i>Nicholay Topin, Stephanie Milani, Fei Fang, Manuela Veloso</i>	
DIFFERENTIALLY PRIVATE AND FAIR DEEP LEARNING: A LAGRANGIAN DUAL APPROACH	9932
<i>Cuong Tran, Ferdinando Fioretto, Pascal Van Hentenryck</i>	
LEARNING ADJUSTMENT SETS FROM OBSERVATIONAL AND LIMITED EXPERIMENTAL DATA	9940
<i>Sofia Triantafillou, Greg Cooper</i>	
*-CFQ: ANALYZING THE SCALABILITY OF MACHINE LEARNING ON A COMPOSITIONAL TASK	9949
<i>Dmitry Tsarkov, Tibor Tihon, Nathan Scales, Nikola Momchev, Danila Sinopalnikov, Nathanael Schärli</i>	
TOWARD ROBUST LONG RANGE POLICY TRANSFER	9958
<i>Wei-Cheng Tseng, Jin-Siang Lin, Yao-Min Feng, Min Sun</i>	
AVOIDING KERNEL FIXED POINTS: COMPUTING WITH ELU AND GELU INFINITE NETWORKS	9967
<i>Russell Tsuchida, Tim Pearce, Chris Van Der Heide, Fred Roosta, Marcus Gallagher</i>	
DEEP FUSION CLUSTERING NETWORK	9978
<i>Wenxuan Tu, Sihang Zhou, Xinwang Liu, Xifeng Guo, Zhiping Cai, En Zhu, Jieren Cheng</i>	
ESCAPED: EFFICIENT SECURE AND PRIVATE DOT PRODUCT FRAMEWORK FOR KERNEL-BASED MACHINE LEARNING ALGORITHMS WITH APPLICATIONS IN HEALTHCARE	9988
<i>Ali Burak Ünal, Mete Akgün, Nico Pfeifer</i>	

EXPECTED ELIGIBILITY TRACES	9997
<i>Hado Van Hasselt, Sephora Madjiheurem, Matteo Hessel, David Silver, André Barreto, Diana Borsa</i>	
CONTINUAL GENERAL CHUNKING PROBLEM AND SYNCMAP	10006
<i>Danilo Vasconcellos Vargas, Toshitake Asabuki</i>	
GATED LINEAR NETWORKS	10015
<i>Joel Veness, Tor Lattimore, David Budden, Avishkar Bhoopchand, Christopher Mattern, Agnieszka Grabska-Barwinska, Eren Sezener, Jianan Wang, Peter Toth, Simon Schmitt, Marcus Hutter</i>	
GRAPHMIX: IMPROVED TRAINING OF GNNS FOR SEMI-SUPERVISED LEARNING	10024
<i>Vikas Verma, Meng Qu, Kenji Kawaguchi, Alex Lamb, Yoshua Bengio, Juho Kannala, Jian Tang</i>	
PID-BASED APPROACH TO ADVERSARIAL ATTACKS	10033
<i>Chen Wan, Biaohua Ye, Fangjun Huang</i>	
NEAREST NEIGHBOR CLASSIFIER EMBEDDED NETWORK FOR ACTIVE LEARNING	10041
<i>Fang Wan, Tianning Yuan, Mengying Fu, Xiangyang Ji, Qingming Huang, Qixiang Ye</i>	
CONTRASTIVE AND GENERATIVE GRAPH CONVOLUTIONAL NETWORKS FOR GRAPH-BASED SEMI-SUPERVISED LEARNING	10049
<i>Sheng Wan, Shirui Pan, Jian Yang, Chen Gong</i>	
APPROXIMATE MULTIPLICATION OF SPARSE MATRICES WITH LIMITED SPACE	10058
<i>Yuanyu Wan, Lijun Zhang</i>	
PROJECTION-FREE ONLINE LEARNING IN DYNAMIC ENVIRONMENTS	10067
<i>Yuanyu Wan, Bo Xue, Lijun Zhang</i>	
PROJECTION-FREE ONLINE LEARNING OVER STRONGLY CONVEX SETS	10076
<i>Yuanyu Wan, Lijun Zhang</i>	
MULTI-VIEW INFORMATION-BOTTLENECK REPRESENTATION LEARNING	10085
<i>Zhibin Wan, Changqing Zhang, Pengfei Zhu, Qinghua Hu</i>	
SEMI-SUPERVISED NODE CLASSIFICATION ON GRAPHS: MARKOV RANDOM FIELDS VS. GRAPH NEURAL NETWORKS	10093
<i>Binghui Wang, Jinyuan Jia, Neil Zhenqiang Gong</i>	
QUANTUM EXPLORATION ALGORITHMS FOR MULTI-ARMED BANDITS	10102
<i>Daochen Wang, Xuchen You, Tongyang Li, Andrew M. Childs</i>	
LEARNING FROM NOISY LABELS WITH COMPLEMENTARY LOSS FUNCTIONS	10111
<i>Deng-Bao Wang, Yong Wen, Lujia Pan, Min-Ling Zhang</i>	
DEBIASING EVALUATIONS THAT ARE BIASED BY EVALUATIONS	10120
<i>Jingyan Wang, Ivan Stelmakh, Yuting Wei, Nihar B. Shah</i>	
ENHANCING UNSUPERVISED VIDEO REPRESENTATION LEARNING BY DECOUPLING THE SCENE AND THE MOTION	10129
<i>Jinpeng Wang, Yuting Gao, Ke Li, Jianguo Hu, Xinyang Jiang, Xiaowei Guo, Rongrong Ji, Xing Sun</i>	

CONSISTENCY REGULARIZATION WITH HIGH-DIMENSIONAL NON-ADVERSARIAL SOURCE-GUIDED PERTURBATION FOR UNSUPERVISED DOMAIN ADAPTATION IN SEGMENTATION	10138
<i>Kaihong Wang, Chenhongyi Yang, Margrit Betke</i>	
EMBEDDING HETEROGENEOUS NETWORKS INTO HYPERBOLIC SPACE WITHOUT META-PATH	10147
<i>Lili Wang, Chongyang Gao, Chenghan Huang, Ruibo Liu, Weicheng Ma, Soroush Vosoughi</i>	
ADVERSARIAL LINEAR CONTEXTUAL BANDITS WITH GRAPH-STRUCTURED SIDE OBSERVATIONS	10156
<i>Lingda Wang, Bingcong Li, Huozhi Zhou, Georgios B. Giannakis, Lav R. Varshney, Zhizhen Zhao</i>	
ADDRESSING CLASS IMBALANCE IN FEDERATED LEARNING	10165
<i>Lixu Wang, Shichao Xu, Xiao Wang, Qi Zhu</i>	
CONTRASTIVE TRANSFORMATION FOR SELF-SUPERVISED CORRESPONDENCE LEARNING	10174
<i>Ning Wang, Wengang Zhou, Houqiang Li</i>	
TACKLING INSTANCE-DEPENDENT LABEL NOISE VIA A UNIVERSAL PROBABILISTIC MODEL	10183
<i>Qizhou Wang, Bo Han, Tongliang Liu, Gang Niu, Jian Yang, Chen Gong</i>	
LEARNING WITH GROUP NOISE	10192
<i>Qizhou Wang, Jiangchao Yao, Chen Gong, Tongliang Liu, Mingming Gong, Hongxia Yang, Bo Han</i>	
ADAPTIVE VERIFIABLE TRAINING USING PAIRWISE CLASS SIMILARITY	10201
<i>Shiqi Wang, Kevin Eykholt, Taesung Lee, Jiyong Jang, Ian Molloy</i>	
ADAPTIVE ALGORITHMS FOR MULTI-ARMED BANDIT WITH COMPOSITE AND ANONYMOUS FEEDBACK	10210
<i>Siwei Wang, Haoyun Wang, Longbo Huang</i>	
HARMONIZED DENSE KNOWLEDGE DISTILLATION TRAINING FOR MULTI-EXIT ARCHITECTURES	10218
<i>Xinglu Wang, Yingming Li</i>	
TIED BLOCK CONVOLUTION: LEANER AND BETTER CNNs WITH SHARED THINNER FILTERS	10227
<i>Xudong Wang, Stella X. Yu</i>	
DEEP RECURRENT BELIEF PROPAGATION NETWORK FOR POMDPS	10236
<i>Yuhui Wang, Xiaoyang Tan</i>	
DATA-FREE KNOWLEDGE DISTILLATION WITH SOFT TARGETED TRANSFER SET SYNTHESIS	10245
<i>Zi Wang</i>	
INCREMENTAL EMBEDDING LEARNING VIA ZERO-SHOT TRANSLATION	10254
<i>Kun Wei, Cheng Deng, Xu Yang, Maosen Li</i>	
GENE REGULATORY NETWORK INFERENCE AS RELAXED GRAPH MATCHING	10263
<i>Deborah Weighill, Marouen Ben Guebila, Camila Lopes-Ramos, Kimberly Glass, John Quackenbush, John Platig, Rebekka Burkholz</i>	

UNIFIED TENSOR FRAMEWORK FOR INCOMPLETE MULTI-VIEW CLUSTERING AND
MISSING-VIEW INFERRING 10273
Jie Wen, Zheng Zhang, Zhao Zhang, Lei Zhu, Lunke Fei, Bob Zhang, Yong Xu

Author Index