



Proceedings of the
Thirtieth AAAI Conference on
Artificial Intelligence and the
Twenty-Eighth Innovative Applications of
Artificial Intelligence Conference

Volume Three

12–17 February 2016
Phoenix, Arizona USA

Palo Alto, California
AAAI Press

Contents

VOLUME III

Main Technical Track (*Continued*)

Machine Learning Methods

Data Poisoning Attacks against Autoregressive Models / 1452

Scott Alfeld, Xiaojin Zhu, Paul Barford

Approximate K-Means++ in Sublinear Time / 1459

Olivier Bachem, Mario Lucic, S. Hamed Hassani, Andreas Krause

Incremental Stochastic Factorization for Online Reinforcement Learning / 1468

André M. S. Barreto, Rafael L. Beirigo, Joelle Pineau, Doina Precup

Increasing the Action Gap: New Operators for Reinforcement Learning / 1476

Marc G. Bellemare, Georg Ostrovski, Arthur Guez, Philip S. Thomas, Rémi Munos

Decoding Hidden Markov Models Faster than Viterbi
Via Online Matrix-Vector (max, +)-Multiplication / 1484

Massimo Cairo, Gabriele Farina, Romeo Rizzi

Maximum Margin Dirichlet Process Mixtures for Clustering / 1491

Gang Chen, Haiying Zhang, Caiming Xiong

Progressive EM for Latent Tree Models and Hierarchical Topic Detection / 1498

Peixian Chen, Nevin L. Zhang, Leonard K. M. Poon, Zhourong Chen

Knowledge Transfer with Interactive Learning of Semantic Relationships / 1505

Jonghyun Choi, Sung Ju Hwang, Leonid Sigal, Larry S. Davis

Robustness of Bayesian Pool-Based Active Learning against Prior Misspecification / 1512

Nguyen Viet Cuong, Nan Ye, Wee Sun Lee

Learning Step Size Controllers for Robust Neural Network Training / 1519

Christian Daniel, Jonathan Taylor, Sebastian Nowozin

Generalised Brown Clustering and Roll-Up Feature Generation / 1526

Leon Derczynski, Sean Chester

Random Composite Forests / 1533

Giulia DeSalvo, Mehryar Mohri

Reconstructing Hidden Permutations Using the Average-Precision (AP) Correlation Statistic / 1540

Lorenzo De Stefani, Alessandro Epasto, Eli Upfal, Fabio Vandin

The Ostomachion Process / 1547

Xuhui Fan, Bin Li, Yi Wang, Yang Wang, Fang Chen

Indexable Probabilistic Matrix Factorization for Maximum Inner Product Search / 1544

Marco Fraccaro, Ulrich Paquet, Ole Winther

- Fast Lasso Algorithm via Selective Coordinate Descent / 1561
Yasuhiro Fujiwara, Yasutoshi Ida, Hiroaki Shiokawa, Sotetsu Iwamura
- Group and Graph Joint Sparsity for Linked Data Classification / 1568
Longwen Gao, Shuigeng Zhou
- Risk Minimization in the Presence of Label Noise / 1575
Wei Gao, Lu Wang, Yu-Feng Li, Zhi-Hua Zhou
- Decentralized Approximate Bayesian Inference for Distributed Sensor Network / 1582
Behnam Gholami, Sejong Yoon, Vladimir Pavlovic
- Assumed Density Filtering Methods for Learning Bayesian Neural Networks / 1589
Soumya Ghosh, Francesco Maria Delle Fave, Jonathan Yedidia
- Extending the Modelling Capacity of Gaussian Conditional Random Fields while Learning Faster / 1596
Jesse Glass, Mohamed Ghalwash, Milan Vukicevic, Zoran Obradovic
- Uncertainty Propagation in Long-Term Structured Regression on Evolving Networks / 1603
Djordje Gligorijevic, Jelena Stojanovic, Zoran Obradovic
- Teaching-to-Learn and Learning-to-Teach for Multi-Label Propagation / 1610
Chen Gong, Dacheng Tao, Jie Yang, Wei Liu
- Discriminative Analysis Dictionary Learning / 1617
Jun Guo, Yanqing Guo, Xiangwei Kong, Man Zhang, Ran He
- Active Learning with Cross-Class Knowledge Transfer / 1624
Yuchen Guo, Guiguang Ding, Yuqi Wang, Xiaoming Jin
- Generalized Emphatic Temporal Difference Learning: Bias-Variance Analysis / 1631
Assaf Hallak, Aviv Tamar, Remi Munos, Shie Mannor
- Multi-Stage Multi-Task Learning with Reduced Rank / 1638
Lei Han, Yu Zhang
- Reduction Techniques for Graph-Based Convex Clustering / 1645
Lei Han, Yu Zhang
- SAND: Semi-Supervised Adaptive Novel Class Detection and Classification over Data Stream / 1652
Ahsanul Haque, Latifur Khan, Michael Baron
- Flattening the Density Gradient for Eliminating Spatial Centrality to Reduce Hubness / 1659
Kazuo Hara, Ikumi Suzuki, Kei Kobayashi, Kenji Fukumizu, Milos Radovanovic
- Discriminative Vanishing Component Analysis / 1666
Chenping Hou, Feiping Nie, Dacheng Tao
- Common and Discriminative Subspace Kernel-Based
 Multiblock Tensor Partial Least Squares Regression / 1673
Ming Hou, Qibin Zhao, Brahim Chaib-draa, Andrzej Cichocki
- Multi-Label Manifold Learning / 1680
Peng Hou, Xin Geng, Min-Ling Zhang
- Optimal Discrete Matrix Completion / 1687
Zhouyuan Huo, Ji Liu, Heng Huang

- Conservativeness of Untied Auto-Encoders / 1694
Daniel Jiwoong Im, Mohamed Ishmael Belghazi, Roland Memisevic
- Infinite Plaid Models for Infinite Bi-Clustering / 1701
Katsuhiko Ishiguro, Issei Sato, Masahiro Nakano, Akisato Kimura, Naonori Ueda
- Improving Predictive State Representations via Gradient Descent / 1709
Nan Jiang, Alex Kulesza, Satinder Singh
- A Probabilistic Approach to Knowledge Translation / 1716
Shangpu Jiang, Daniel Lowd, Dejing Dou
- The $\ell_{2,1}$ -Norm Stacked Robust Autoencoders for Domain Adaptation / 1723
Wenhao Jiang, Hongchang Gao, Fu-lai Chung, Heng Huang
- Wishart Mechanism for Differentially Private Principal Components Analysis / 1730
Wuxuan Jiang, Cong Xie, Zhihua Zhang
- Deep Learning with S-Shaped Rectified Linear Activation Units / 1737
Xiaojie Jin, Chunyan Xu, Jiashi Feng, Yunchao Wei, Junjun Xiong, Shuicheng Yan
- Delay-Tolerant Online Convex Optimization: Unified Analysis
 and Adaptive-Gradient Algorithms / 1744
Pooria Joulani, András György, Csaba Szepesvári
- Shakeout: A New Regularized Deep Neural Network Training Scheme / 1751
Guoliang Kang, Jun Li, Dacheng Tao
- Bounded Optimal Exploration in MDP / 1758
Kenji Kawaguchi
- Uncorrelated Group LASSO / 1765
Deguang Kong, Ji Liu, Bo Liu, Xuan Bao
- Learning Future Classifiers without Additional Data / 1772
Atsutoshi Kumagai, Tomoharu Iwata
- Compressed Conditional Mean Embeddings for Model-Based Reinforcement Learning / 1779
Guy Lever, John Shawe-Taylor, Ronnie Stafford, Csaba Szepesvári
- Preconditioned Stochastic Gradient Langevin Dynamics for Deep Neural Networks / 1788
Chunyuan Li, Changyou Chen, David Carlson, Lawrence Carin
- High-Order Stochastic Gradient Thermostats for Bayesian Learning of Deep Models / 1795
Chunyuan Li, Changyou Chen, Kai Fan, Lawrence Carin
- Multi-Objective Self-Paced Learning / 1802
Hao Li, Maoguo Gong, Deyu Meng, Qiguang Miao
- Scalable Sequential Spectral Clustering / 1809
Yeqing Li, Junzhou Huang, Wei Liu
- Towards Safe Semi-Supervised Learning for Multivariate Performance Measures / 1816
Yu-Feng Li, James T. Kwok, Zhi-Hua Zhou
- Accelerating Random Kaczmarz Algorithm Based on Clustering Information / 1823
Yujun Li, Kaichun Mo, Haishan Ye

- Fast and Accurate Refined Nyström Based Kernel SVM / 1830
Zhe Li, Tianbao Yang, Lijun Zhang, Rong Jin
- How Important Is Weight Symmetry in Backpropagation? / 1837
Qianli Liao, Joel Z. Leibo, Tomaso Poggio
- Re-Active Learning: Active Learning with Relabeling / 1845
Christopher H. Lin, Mausam, Daniel S. Weld
- Interaction Point Processes via Infinite Branching Model / 1853
Peng Lin, Bang Zhang, Ting Guo, Yang Wang, Fang Chen
- Gaussian Process Planning with Lipschitz Continuous Reward Functions:
Towards Unifying Bayesian Optimization, Active Learning, and Beyond / 1860
Chun Kai Ling, Kian Hsiang Low, Patrick Jaillet
- Online ARIMA Algorithms for Time Series Prediction / 1867
Chenghao Liu, Steven C. H. Hoi, Peilin Zhao, Jianling Sun
- Consensus Guided Unsupervised Feature Selection / 1874
Hongfu Liu, Ming Shao, Yun Fu
- Sparse Perceptron Decision Tree for Millions of Dimensions / 1881
Weiwei Liu, Ivor W. Tsang
- Multiple Kernel k-Means Clustering with Matrix-Induced Regularization / 1888
Xinwang Liu, Yong Dou, Jianping Yin, Lei Wang, En Zhu
- Finding One's Best Crowd: Online Learning by Exploiting Source Similarity / 1895
Yang Liu, Mingyan Liu
- Learning FRAME Models Using CNN Filters / 1902
Yang Lu, Song-Chun Zhu, Ying Nian Wu
- Sparse Latent Space Policy Search / 1911
Kevin Sebastian Luck, Joni Pajarinen, Erik Berger, Ville Kyrki, Heni Ben Amor
- Expected Tensor Decomposition with Stochastic Gradient Descent / 1919
Takanori Maehara, Kohei Hayashi, Ken-ichi Kawarabayashi
- Offline Evaluation of Online Reinforcement Learning Algorithms / 1926
Travis Mandel, Yun-En Liu, Emma Brunskill, Zoran Popović
- Reinforcement Learning with Parameterized Actions / 1934
Warwick Masson, Praveesh Ranchod, George Konidaris
- Fixed-Rank Supervised Metric Learning on Riemannian Manifold / 1941
Yadong Mu
- All-in Text: Learning Document, Label, and Word Representations Jointly / 1948
Jinseok Nam, Eneldo Loza Mencla, Johannes Fürnkranz
- Holographic Embeddings of Knowledge Graphs / 1955
Maximilian Nickel, Lorenzo Rosasco, Tomaso Poggio
- New ℓ_1 -Norm Relaxations and Optimizations for Graph Clustering / 1962
Feiping Nie, Hua Wang, Cheng Deng, Xinbo Gao, Xuelong Li, Heng Huang

- The Constrained Laplacian Rank Algorithm for Graph-Based Clustering / 1969
Feiping Nie, Xiaoqian Wang, Michael I. Jordan, Heng Huang
- Efficient PAC-Optimal Exploration in Concurrent, Continuous State MDPs with Delayed Updates / 1977
Jason Pazis, Ronald Parr
- Viral Clustering: A Robust Method to Extract Structures in Heterogeneous Datasets / 1986
Vahan Petrosyan, Alexandre Proutiere
- Inverse Reinforcement Learning through Policy Gradient Minimization / 1993
Matteo Pirodda, Marcello Restelli
- Scaling Simultaneous Optimistic Optimization for High-Dimensional Non-Convex Functions with Low Effective Dimensions / 2000
Hong Qian, Yang Yu
- Selecting Near-Optimal Learners via Incremental Data Allocation / 2007
Ashish Sabharwal, Horst Samulowitz, Gerald Tesaro
- Scalable Algorithms for Tractable Schatten Quasi-Norm Minimization / 2016
Fanhua Shang, Yuanyuan Liu, James Cheng
- Spectral Bisection Tree Guided Deep Adaptive Exemplar Autoencoder for Unsupervised Domain Adaptation / 2023
Ming Shao, Zhengming Ding, Handong Zhao, Yun Fu
- Metric Learning for Ordinal Data / 2030
Yuan Shi, Wenzhe Li, Fei Sha
- Noisy Submodular Maximization via Adaptive Sampling with Applications to Crowdsourced Image Collection Summarization / 2037
Adish Singla, Sebastian Tschiatschek, Andreas Krause
- Bayesian Matrix Completion via Adaptive Relaxed Spectral Regularization / 2044
Yang Song, Jun Zhu
- Marginalized Continuous Time Bayesian Networks for Network Reconstruction from Incomplete Observations / 2051
Lukas Studer, Loïc Paulevé, Christoph Zechner, Matthias Reumann, María Rodríguez Martínez, Heinz Koepl
- Return of Frustratingly Easy Domain Adaptation / 2058
Baochen Sun, Jiashi Feng, Kate Saenko
- On the Depth of Deep Neural Networks: A Theoretical View / 2066
Shizhao Sun, Wei Chen, Liwei Wang, Xiaoguang Liu, Tie-Yan Liu
- Linear-Time Learning on Distributions with Approximate Kernel Embeddings / 2073
Dougal J. Sutherland, Junier B. Oliva, Barnabás Póczos, Jeff Schneider
- Learning Sparse Confidence-Weighted Classifier on Very High Dimensional Data / 2080
Mingkui Tan, Yan Yan, Li Wang, Anton Van Den Hengel, Ivor W. Tsang, Qinfeng (Javen) Shi
- Algorithms for Differentially Private Multi-Armed Bandits / 2087
Aristide C. Y. Tossou, Christos Dimitrakakis
- Deep Reinforcement Learning with Double Q-Learning / 2094
Hado van Hasselt, Arthur Guez, David Silver

- Online Instrumental Variable Regression with Applications
to Online Linear System Identification / 2101
Arun Venkatraman, Wen Sun, Martial Hebert, J. Andrew Bagnell, Byron Boots
- The Hidden Convexity of Spectral Clustering / 2108
James Voss, Mikhail Belkin, Luis Rademacher
- Multitask Generalized Eigenvalue Program / 2115
Boyu Wang, Joelle Pineau, Borja Balle
- Product Grassmann Manifold Representation and Its LRR Models / 2122
Boyue Wang, Yongli Hu, Junbin Gao, Yanfeng Sun, Baocai Yin
- Text Classification with Heterogeneous Information Network Kernels / 2130
Chenguang Wang, Yangqiu Song, Haoran Li, Ming Zhang, Jiawei Han
- Semi-Supervised Dictionary Learning via Structural Sparse Preserving / 2137
Di Wang, Xiaoqin Zhang, Mingyu Fan, Xiuzi Ye
- Relational Knowledge Transfer for Zero-Shot Learning / 2145
Donghui Wang, Yanan Li, Yuetan Lin, Yueting Zhuang
- Optimizing Multivariate Performance Measures from Multi-View Data / 2152
Jim Jing-Yan Wang, Ivor Wai-Hung Tsang, Xin Gao
- An Efficient Time Series Subsequence Pattern Mining and
Prediction Framework with an Application to Respiratory Motion Prediction / 2159
Shouyi Wang, Kin Ming Kam, Cao Xiao, Stephen Bowen, W. Chaovalitwongse
- Co-Regularized PLSA for Multi-Modal Learning / 2166
Xin Wang, Ming-Ching Chang, Yiming Ying, Siwei Lyu
- Toward a Better Understanding of Deep Neural Network Based
Acoustic Modelling: An Empirical Investigation / 2173
Xingfu Wang, Lin Wang, Jing Chen, Litao Wu
- Noise-Adaptive Margin-Based Active Learning and Lower Bounds
under Tsybakov Noise Condition / 2180
Yining Wang, Aarti Singh
- Learning by Transferring from Unsupervised Universal Sources / 2187
Yu-Xiong Wang, Martial Hebert
- Learning Deep ℓ_0 Encoders / 2194
Zhangyang Wang, Qing Ling, Thomas S. Huang
- Adaptive Normalized Risk-Averting Training for Deep Neural Networks / 2201
Zhiguang Wang, Tim Oates, James Lo
- Nonlinear Feature Extraction with Max-Margin Data Shifting / 2208
Jianqiao Wangni, Ning Chen
- Unsupervised Feature Selection on Networks: A Generative View / 2215
Xiaokai Wei, Bokai Cao, Philip S. Yu
- Model-Free Preference-Based Reinforcement Learning / 2222
Christian Wirth, Johannes Fürnkranz, Gerhard Neumann

- Constrained Submodular Minimization for Missing Labels and Class Imbalance in Multi-Label Learning / 2229
Baoyuan Wu, Siwei Lyu, Bernard Ghanem
- Representing Sets of Instances for Visual Recognition / 2237
Jianxin Wu, Bin-Bin Gao, Guoqing Liu
- Robust Semi-Supervised Learning through Label Aggregation / 2244
Yan Yan, Zhongwen Xu, Ivor W. Tsang, Guodong Long, Yi Yang
- Analysis-Synthesis Dictionary Learning for Universality-Particularity Representation Based Classification / 2251
Meng Yang, Weiyang Liu, Weixin Luo, Linlin Shen
- Efficient Average Reward Reinforcement Learning Using Constant Shifting Values / 2258
Shangdong Yang, Yang Gao, Bo An, Hao Wang, Xingguo Chen
- Learning Continuous-Time Bayesian Networks in Relational Domains: A Non-Parametric Approach / 2265
Shuo Yang, Tushar Khot, Kristian Kersting, Sriraam Natarajan
- Instance Specific Metric Subspace Learning: A Bayesian Approach / 2272
Han-Jia Ye, De-Chuan Zhan, Yuan Jiang
- Scalable Completion of Nonnegative Matrices with the Separable Structure / 2279
Xiyu Yu, Wei Bian, Dacheng Tao
- Derivative-Free Optimization via Classification / 2286
Yang Yu, Hong Qian, Yi-Qi Hu
- On Order-Constrained Transitive Distance Clustering / 2293
Zhiding Yu, Weiyang Liu, Wenbo Liu, Yingzhen Yang, Ming Li, B. V. K. Vijaya Kumar
- A Proximal Alternating Direction Method for Semi-Definite Rank Minimization / 2300
Ganzhao Yuan, Bernard Ghanem
- Learning Expected Hitting Time Distance / 2309
De-Chuan Zhan, Peng Hu, Zui Chu, Zhi-Hua Zhou
- Stochastic Optimization for Kernel PCA / 2316
Lijun Zhang, Tianbao Yang, Jinfeng Yi, Rong Jin, Zhi-Hua Zhou
- Asynchronous Distributed Semi-Stochastic Gradient Optimization / 2323
Ruiliang Zhang, Shuai Zheng, James T. Kwok
- An Alternating Proximal Splitting Method with Global Convergence for Nonconvex Structured Sparsity Optimization / 2330
Shubao Zhang, Hui Qian, Xiaojin Gong
- Accelerated Sparse Linear Regression via Random Projection / 2337
Weizhong Zhang, Lijun Zhang, Rong Jin, Deng Cai, Xiaofei He
- Large-Scale Graph-Based Semi-Supervised Learning via Tree Laplacian Solver / 2344
Yan-Ming Zhang, Xu-Yao Zhang, Xiao-Tong Yuan, Cheng-Lin Liu
- Near-Optimal Active Learning of Multi-Output Gaussian Processes / 2351
Yehong Zhang, Trong Nghia Hoang, Kian Hsiang Low, Mohan Kankanhalli

- Multi-Domain Active Learning for Recommendation / 2358
Zihan Zhang, Xiaoming Jin, Lianghao Li, Guiguang Ding, Qiang Yang
- On the Differential Privacy of Bayesian Inference / 2365
Zuhe Zhang, Benjamin I. P. Rubinstein, Christos Dimitrakakis
- A Scalable and Extensible Framework for Superposition-Structured Models / 2372
Shenjian Zhao, Cong Xie, Zhihua Zhang
- Fast Asynchronous Parallel Stochastic Gradient Descent:
A Lock-Free Approach with Convergence Guarantee / 2379
Shen-Yi Zhao, Wu-Jun Li
- DinTucker: Scaling Up Gaussian Process Models on Large Multidimensional Arrays / 2386
Shandian Zhe, Yuan Qi, Youngja Park, Zenglin Xu, Ian Molloy, Suresh Chari
- Fast Nonsmooth Regularized Risk Minimization with Continuation / 2393
Shuai Zheng, Ruiliang Zhang, James T. Kwok
- Transfer Learning for Cross-Language Text Categorization
through Active Correspondences Construction / 2400
Joey Tianyi Zhou, Sinno Jialin Pan, Ivor W. Tsang, Shen-Shyang Ho
- Veto-Consensus Multiple Kernel Learning / 2407
Yuxun Zhou, Ninghang Hu, Costas J. Spanos
- Deep Hashing Network for Efficient Similarity Retrieval / 2415
Han Zhu, Mingsheng Long, Jianmin Wang, Yue Cao
- Coupled Dictionary Learning for Unsupervised Feature Selection / 2422
Pengfei Zhu, Qinghua Hu, Changqing Zhang, Wangmeng Zuo
- Stochastic Parallel Block Coordinate Descent for Large-Scale Saddle Point Problems / 2429
Zhanxing Zhu, Amos J. Storkey