Relative Fisher Information and Natural Gradient for Learning Large Modular Models
Ke Sun (King Abdullah University of Science and Technology) · Frank Nielsen (École Polytechnique)

Priv'IT: Private and Sample Efficient Identity Testing
Bryan Cai (MIT) · Constantinos Daskalakis (MIT) · Gautam Kamath (MIT)

Being Robust (in High-Dimensions) Can Be Practical
Ilias Diakonikolas (USC) · Gautam Kamath (MIT) · Daniel Kane (UCSD) · Jerry Li (MIT) · Ankur Moitra () · Alistair Stewart (USC)

Unifying task specification in reinforcement learning
Martha White (Indiana University)

Fractional Langevin Monte Carlo: Exploring Levy Driven Stochastic Differential Equations for MCMC
Umut Simsekli (Telecom ParisTech)

Lost Relatives of the Gumbel Trick
Matej Balog (University of Cambridge) · Nilesh Tripuraneni () · Zoubin Ghahramani (University of Cambridge) · Adrian Weller (University of Cambridge)

Learning the Structure of Generative Models without Labeled Data
Stephen Bach (Stanford University) · Bryan He (Stanford University) · Alexander J Ratner (Stanford University) · Christopher Re (Stanford)

Deep Tensor Convolution on Multicores
David Budden (MIT) · Alexander Matveev (MIT) · Shibani Santurkar (MIT) · Shraman Chaudhuri (MIT) · Nir Shavit (MIT)

Beyond Filters: Compact Feature Map for Portable Deep Model
Yunhe Wang (Peking University) · Chang Xu (University of Sydney) · Chao Xu (Peking University) · Dacheng Tao ()

Tight Bounds for Approximate Carathéodory and Beyond
Vahab Mirrokni (Google Research) · Renato Leme (Google Research) · Adrian Vladu (MIT) · Sam Wong (UC Berkeley)

Fast k-Nearest Neighbour Search via Prioritized DCI
Ke Li (UC Berkeley) · Jitendra Malik ()

An Adaptive Test of Independence with Analytic Kernel Embeddings
Wittawat Jitkrittum (UCL) · Zoltan Szabo (École Polytechnique) · Arthur Gretton (Gatsby)

Deep Transfer Learning with Joint Adaptation Networks
Mingsheng Long (Tsinghua University) · Han Zhu (Tsinghua University) · Jianmin Wang (Tsinghua University) · Michael Jordan ()

Robust Probabilistic Modeling with Bayesian Data Reweighting
Yixin Wang (Columbia University) · Alp Kucukelbir (Columbia University) · David Blei (Columbia University)

Distributed and Provably Good Seedings for k-Means in Constant Rounds
Olivier Bachem (ETH Zurich) · Mario Lucic (ETH Zurich) · Andreas Krause (ETH Zurich)

Toward Efficient and Accurate Covariance Matrix Estimation on Compressed Data
XIXIAN CHEN (The Chinese University of Hong Kong) · Irwin King (CUHK) · Michael Lyu (The Chinese University of Hong Kong)

Combined Group and Exclusive Sparsity for Deep Neural Networks
jaehong yoon (UNIST) · Sung Hwang ()

Robust Guarantees of Stochastic Greedy Algorithms
Yaron Singer (Harvard) · Avinatan Hassidim (Bar Ilan University)

Analysis and Optimization of Graph Decompositions by Lifted Multicuts
Andrea Hornakova (Max Planck Institute for Informatics) · Jan-Hendrik Lange (MPI for Informatics) · Bjoern Andres (MPI for Informatics)

GSOS: Gauss-Seidel Operator Splitting Algorithm for Multi-Term Nonsmooth Convex Composite Optimization
Li Shen (School of Mathematics, South China University of Technology) · Wei Liu (Tencent AI Lab) · GanZhao Yuan () · Shiqian Ma (The Chinese University of Hong Kong)

Curiosity-driven Exploration by Self-supervised Prediction
Deepak Pathak (UC Berkeley) · Pulkit Agrawal () · Alexei Efros (UC Berkeley) · Trevor Darrell (University of California at Berkeley)

Uncertainty Assessment and False Discovery Rate Control in High-Dimensional Granger Causal Inference
Aditya Chaudhry (University of Virginia) · Pan Xu (University of Virginia) · Quanquan Gu (University of Virginia)

Consistent On-Line Off-Policy Evaluation
Assaf Hallak (Technion) · Shie Mannor (Technion)

Coresets for Vector Summarization with Applications to Network Graphs
Dan Feldman () · Sedat Ozer (MIT) · Daniela Rus ()

Oracle Complexity of Second-Order Methods for Finite-Sum Problems
Yossi Arjevani (Weizmann Institute of Science) · Ohad Shamir (Weizmann Institute of Science)

Active Learning for Accurate Estimation of Linear Models
Carlos Riquelme Ruiz (Stanford University) · Mohammad Ghavamzadeh (Adobe) · Alessandro Lazaric ()

Multiple Clustering Views from Multiple Uncertain Experts
Yale Chang (Northeastern University) · Junxiang Chen (Northeastern University) · Michael Cho (Harvard Medical School) · Peter Castaldi (Harvard Medical School) · Edwin Silverman (Harvard Medical School) · Jennifer G Dy (Northeastern University)

Doubly Accelerated Methods for Faster CCA and Generalized Eigendecomposition
Zeyuan Allen-Zhu (Institute for Advanced Study) · Yuanzhi Li (Princeton University)

Sketched Ridge Regression: Optimization Perspective, Statistical Perspective, and Model Averaging
Shusen Wang (UC Berkeley) · Alex Gittens (RPI) · Michael Mahoney (UC Berkeley)

When can Multi-Site Datasets be Pooled for Regression? Hypothesis Tests, ℓ_2ℓ2-consistency and Neuroscience Applications
Hao Zhou (University of Wisconsin - Madison) · Yilin Zhang () · Vamsi Ithapu (Univresity of Wisconsin Madiso) · Sterling Johnson (UW Madison) · Grace Wahba () · Vikas Singh ()

Learning Deep Architectures via Generalized Whitened Neural Networks
Ping Luo (The Chinese University of Hong Kong)

How close are the eigenvectors and eigenvalues of the sample and actual covariance matrices?
Andreas Loukas (EPFL)

SplitNet: Learning to Semantically Split Deep Networks for Parameter Reduction and Model Parallelization
Juyong Kim (Seoul National University) · Yookoon Park (Seoul National University) · Gunhee Kim (Seoul National University) · Sung Hwang ()

Uncorrelation and Evenness: A New Diversity-Promoting Regularizer
Pengtao Xie (Carnegie Mellon University) · Aarti Singh () · Eric Xing (Carnegie Mellon University)

Follow the Compressed Leader: Even Faster Online Learning of Eigenvectors
Zeyuan Allen-Zhu (Institute for Advanced Study) · Yuanzhi Li (Princeton University)

Faster Principal Component Regression via Optimal Polynomial Approximation to Matrix sgn(x)
Zeyuan Allen-Zhu (Institute for Advanced Study) · Yuanzhi Li (Princeton University)

Deep Spectral Clutering Learning
Marc Law (University of Toronto) · Raquel Urtasun (University of Toronto) · Zemel Rich (University of Toronto)

Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks
Chelsea Finn (UC Berkeley) · Pieter Abbeel (OpenAI / Berkeley) · Sergey Levine (Berkeley)

Learning to Discover Cross-Domain Relations with Generative Adversarial Networks
Taeksoo Kim () · (None) · Hyunsoo Kim (SK T-Brain) · Jungkwon Lee (SK T-Brain) · Joseph Lim (Univ. of Southern California) · Jiwon Kim (SK T-Brain)

Dynamic Word Embeddings via Skip-Gram Filtering
Stephan Mandt (Disney Research) · Robert Bamler (Disney Research Pittsburgh)

Image-to-Markup Generation with Coarse-to-Fine Attention
Yuntian Deng (Harvard University) · Anssi Kanervisto (University of Eastern Finland) · Jeffrey Ling (Harvard University) · Alexander Rush (Harvard University)

Cosine Similarity Constrained Latent Space Models
Pengtao Xie (Carnegie Mellon University) · Yuntian Deng (Harvard University) · Yi Zhou (Syracuse University) · Abhimanu Kumar (Groupon Inc.) · Yaoliang Yu (University of Waterloo) · James Zou (Stanford) · Eric Xing (Carnegie Mellon University)

Orthogonalized ALS: A Theoretically Principled Tensor Decomposition Algorithm for Practical Use
Vatsal Sharan (Stanford University) · Gregory Valiant (Stanford University)

Regret Minimization in Behaviorally-Constrained Zero-Sum Games
Gabriele Farina (Carnegie Mellon University) · Christian Kroer (Carnegie Mellon University) · Tuomas Sandholm (Carnegie Mellon University)

Breaking Locality Accelerates Block Gauss-Seidel
Stephen Tu (UC Berkeley) · Shivaram Venkataraman (UC Berkeley) · Ashia Wilson (UC Berkeley) · Alex Gittens (UC Berkeley) · Michael Jordan () · Benjamin Recht (Berkeley)

Learning to Aggregate Ordinal Labels by Maximizing Separating Width
Guangyong Chen (The Chinese University of Hong Kong) · Shengyu Zhang (CUHK) · Di Lin (Shenzhen University) · Hui Huang (Shenzhen University) · Pheng Ann Heng (The Chinese University of Hong Kong)

Composing Tree Graphical Models with Persistent Homology Features for Clustering Mixed-Type Data
XIUYAN NI (THE GRADUATE CENTER, CUNY) · Novi Quadrianto (University of Sussex and National Research University Higher School of Economics) · Yusu Wang (Ohio State University) · Chao Chen (CUNY Queens College)

Stochastic Convex Optimization: Faster Local Growth Implies Faster Global Convergence
Yi Xu () · Qihang Lin (Univ Iowa) · Tianbao Yang ()

Scalable Multi-Class Gaussian Process Classification using Expectation Propagation
Carlos Villacampa-Calvo (Universidad Autónoma de Madrid) · Daniel Hernandez-Lobato (Universidad Autonoma de Madrid)

Canopy --- Fast Sampling with Cover Trees
Manzil Zaheer (CMU) · Satwik Kottur (Carnegie Mellon University) · Amr Ahmed (Google) · Jose Moura (CMU) · Alex Smola (Amazon)

Magnetic Hamiltonian Monte Carlo
Nilesh Tripuraneni () · Mark Rowland (University of Cambridge) · Zoubin Ghahramani (University of Cambridge) · Rich Turner (University of Cambridge)

Lazifying Conditional Gradient Algorithms
Gábor Braun () · Sebastian Pokutta (Georgia Tech) · Daniel Zink ()

Conditional Accelerated Lazy Stochastic Gradient Descent
Guanghui Lan () · Sebastian Pokutta (Georgia Tech) · Yi Zhou () · Daniel Zink ()

A Richer Theory of Convex Constrained Optimization with Reduced Projections and Improved Rates
Tianbao Yang () · Qihang Lin (Univ Iowa) · Lijun Zhang (Nanjing University)

A Semismooth Newton Method for Fast, Generic Convex Programming
Alnur Ali (Carnegie Mellon University) · Eric Wong (Carnegie Mellon University) · Zico Kolter (Carnegie Mellon University)

Sequence Modeling via Segmentations
Chong Wang (Microsoft Research) · Yining Wang (CMU) · Po-Sen Huang (Microsoft Research) · Abdelrahman Mohammad (Microsoft) · Dengyong Zhou (Microsoft Research) · Li Deng (Microsoft Research)

Evaluating Bayesian Models with Posterior Dispersion Indices
Alp Kucukelbir (Columbia University) · David Blei (Columbia University)

State-Frequency Memory Recurrent Neural Networks
Hao Hu (University of Central Florida) · Guo-Jun Qi (University of Central Florida)

Kernelized Tensor Factorization Machines with Applications to Neuroimaging
Lifang He (University of Illinios at Chicago/Shenzhen University) · Chun-Ta Lu (University of Illinois at Chicago) · Guixiang Ma () · Shen Wang (University of Illinios at Chicago) · Linlin Shen () · Philip Yu () · Ann Ragin (Northwestern University)

Re-revisiting Learning on Hypergraphs: Confidence Interval and Subgradient Method
Chenzi Zhang (HKU) · Shuguang Hu (University of Hong Kong) · Zhihao Tang (University of Hong Kong) · Hubert Chan (University of Hong Kong)

Self-Paced Cotraining
Fan Ma (Xian Jiaotong University) · Deyu Meng () · Qi Xie () · Zina Li () · Xuanyi Dong (University of Technology Sydney)

ChoiceRank: Identifying Preferences from Node Traffic in Networks
Lucas Maystre (EPFL) · Matthias Grossglauser (EPFL)

Unsupervised Learning by Predicting Noise
Piotr Bojanowski (Facebook) · Armand Joulin (Facebook)

Guarantees for Greedy Maximization of Non-submodular Functions with Applications
Yatao Bian (ETH Zurich) · Joachim Buhmann () · Andreas Krause (ETH Zurich) · Sebastian Tschiatschek (ETH)

Nonnegative Matrix Factorization for Time Series Recovery From a Few Temporal Aggregates
Jiali Mei (Université Paris-Sud & EDF Lab) · Yohann De Castro (LMO) · Yannig Goude (EDF Lab Paris-Saclay) · Georges Hébrail (EDF Lab Paris-Saclay)

Uniform Deviation Bounds for Unbounded Loss Functions like k-Means
Olivier Bachem (ETH Zurich) · Mario Lucic (ETH Zurich) · Hamed Hassani (ETH Zurich) · Andreas Krause (ETH Zurich)

Sliced Wasserstein Kernel for Persistence Diagrams
Mathieu Carrière (Inria Saclay) · Marco Cuturi (ENSAE / CREST) · Steve Oudot ()

Dual Iterative Hard Thresholding: From Non-convex Sparse Minimization to Non-smooth Concave Maximization
Bo Liu (Rutgers) · Xiaotong Yuan (Nanjing University of Information Science & Technology) · Lezi Wang (Rutgers) · Qingshan Liu () · Dimitris Metaxas (Rutgers)

Measuring Sample Quality with Kernels
Jackson Gorham (STANFORD) · Lester Mackey (Microsoft Research)

Coherence Pursuit: Fast, Simple, and Robust Subspace Recovery
Mostafa Rahmani (University of Central Florida) · George Atia (University of Central Florida)

Bidirectional learning for time-series models with hidden units
Takayuki Osogami (IBM Research - Tokyo) · Hiroshi Kajino () · Taro Sekiyama ()

Neural Message Passing for Quantum Chemistry
Justin Gilmer (Google Brain) · Samuel Schoenholz (Google Brain) · Patrick Riley (Google) · Oriol Vinyals (DeepMind) · George Dahl (Google Brain)

Stochastic modified equations and adaptive stochastic gradient algorithms
*Qianxiao Li (Institute of High Performance Computing, A*STAR) · Cheng Tai (Peking University) · Weinan E (Princeton University)*

Learning Stable Stochastic Nonlinear Dynamical Systems
Jonas Umlauft (Technical University of Munich) · Sandra Hirche (Technical University of Munich)

Post-Inference Prior Swapping
William Neiswanger (CMU) · Eric Xing (Carnegie Mellon University)

Online Learning with Local Permutations and Delayed Feedback
Liran Szlak (Weizmann Institute of Science) · Ohad Shamir (Weizmann Institute of Science)

Deep Value Networks Learn to Evaluate and Iteratively Refine Structured Outputs
Michael Gygli (Gifs.com) · Mohammad Norouzi (Google) · Anelia Angelova (Google Brain)

Delta Networks for Optimized Recurrent Network Computation
Daniel Neil (Institute of Neuroinformatics) · Jun Lee (Samsung Advanced Institute of Technology) · Tobi Delbruck (Institute of Neuroinformatics) · Shih-Chii Liu (Institute of Neuroinformatics)

Cognitive Psychology for Deep Neural Networks: A Shape Bias Case Study
Samual Ritter (DeepMind) · David GT Barrett (DeepMind) · Adam Santoro (DeepMind) · Matthew Botvinick (DeepMind)

Spherical Structured Feature Maps for Kernel Approximation
Yueming LYU (city university of hong kong)

Enumerating distinct decision trees
Salvatore Ruggieri (University of Pisa)

Dropout Inference in Bayesian Neural Networks with Alpha-divergences
Yingzhen Li (University of Cambridge) · Yarin Gal (University of Cambridge) · Rich Turner (University of Cambridge)

Convexified Convolutional Neural Networks
Yuchen Zhang (Stanford) · Percy Liang (Stanford University) · Martin Wainwright (University of California at Berkeley)

Automatic Discovery of the Statistical Types of Variables in a Dataset
Isabel Valera () · Zoubin Ghahramani (University of Cambridge)

FeUdal Networks for Hierarchical Reinforcement Learning
Alexander Vezhnevets (DeepMind) · Simon Osindero (DeepMind) · Tom Schaul (DeepMind) · Nicolas Heess (Google DeepMind) · Max Jaderberg (DeepMind) · David Silver (Google DeepMind) · Koray Kavukcuoglu (DeepMind)

Learning Hawkes Processes from Short Doubly-Censored Event Sequences
Hongteng Xu (Georgia Institute of Technolog) · Dixin Luo (University of Toronto) · Hongyuan Zha (Georgia Institute of Technology)

Real-Time Adaptive Image Compression
Oren Rippel (WaveOne, Inc.) · Lubomir Bourdev (WaveOne, Inc.)

Multivariate Kernel Density Estimation: Optimal Uniform Rates
Hanxi Jiang (Google)

Adaptive Multiple-Arm Identification
Jiecao Chen (Indiana University Bloomington) · Xi Chen (New York University) · Qin Zhang (Indiana University Bloomington) · Yuan Zhou (Indiana University Bloomington)

Accelerated Stochastic Gradient Expectation-Maximization Algorithm
Rongda Zhu (UIUC) · Lingxiao Wang (University of Virginia) · Chengxiang Zhai (University of Illinois at Urbana-Champaign) · Quanquan Gu (University of Virginia)

Modular Multitask Reinforcement Learning with Policy Sketches
Jacob Andreas (UC Berkeley) · Sergey Levine (Berkeley) · Dan Klein (UC Berkeley)

Accelerating Eulerian Fluid Simulation With Convolutional Networks
Jonathan Tompson (Google Inc.) · Kristofer Schlachter (New York University) · Pablo Sprechmann (NYU) · Ken Perlin (New York University)

An Analytical Formula of Population Gradient for two-layered ReLU network and its Applications in Convergence and Critical Point Analysis
Yuandong Tian (Facebook)

Partitioned Tensor Factorizations for Learning Mixed Membership Models
Zilong Tan (Duke University) · Sayan Mukherjee (Duke University)

Density Level Set Estimation on Manifolds with DBSCAN
Hanxi Jiang (Google)

Efficient Nonmypoic Active Search
Shali Jiang (Washington University in St. Louis) · Luiz Gustavo Malkomes Muniz (Washington University in St. Louis) · Geoff Converse (Simpson College) · Alyssa Shofner (University of South Carolina) · Benjamin Moseley (Washington University in St. Louis) · Roman Garnett (Washington University in St. Louis)

High Dimensional Bayesian Optimization with Elastic Gaussian Process
Santu Rana (Deakin University) · Cheng Li (Deakin University) · Vu Nguyen (Deakin University) · Sunil Gupta (Deakin University) · Svetha Venkatesh (Deakin University)

Leveraging Node Attributes for Incomplete Relational Data
He Zhao (FIT, Monash University) · Lan Du (Monash University) · Wray Buntine (Monash University)

Tensor Decomposition with Smoothness
Masaaki Imaizumi (Institute of Statistical Mathematics) · Kohei Hayashi (AIST)

Efficient Online Bandit Multiclass Learning with $\tilde{O}(\sqrt{T})$ **Regret**

Alina Beygelzimer (Yahoo Research) · Francesco Orabona (Stony Brook University) · Chicheng Zhang (UCSD)

Variational Boosting: Iteratively Refining Posterior Approximations
Andrew Miller (Harvard) · Nick Foti (University of Washington) · Ryan Adams (Google Brain and Princeton University)

Communication-efficient Algorithms for Distributed Stochastic Principal Component Analysis
Dan Garber (TTIC) · Ohad Shamir (Weizmann Institute of Science) · Nati Srebro (Toyota Technological Institute at Chicago)

Tensor Decomposition via Simultaneously Power iteration
Poan Wang (Academia Sinica) · Chi-Jen Lu (Academia Sinica)

Joint Dimensionality Reduction and Metric Learning: A Geometric Take
Mehrtash Harandi (Data61) · Mathieu Salzmann (EPFL) · Richard Hartley (Australian National University)

Adaptive Sampling Probabilities for Non-Smooth Optimization
Hongseok Namkoong (Stanford University) · Aman Sinha (Stanford University) · Steven Yadlowsky (Stanford University) · John Duchi (Stanford University)

Sub-sampled Cubic Regularization for Non-convex Optimization
Jonas Kohler (ETH Zurich) · Aurelien Lucchi (ETH)

Asynchronous Stochastic Gradient Descent with Delay Compensation
Shuxin Zheng (University of Science and Technology of China) · Qi Meng (Peking University) · Taifeng Wang () · Wei Chen (Microsoft Research) · Tie-Yan Liu (Microsoft)

Preferential Bayesian Optmization
Javier González (Amazon) · Zhenwen Dai (Amazon.com) · Andreas Damianou (Amazon.com) · Neil Lawrence (Amazon.com)

Averaged-DQN: Variance Reduction and Stabilization for Deep Reinforcement Learning
Oron Anschel (Technion) · Nir Baram (Technion) · Nahum Shimkin (Technion)

meProp: Minimal Effort Back Propagation for Accelerated Deep Learning
Xu SUN (Peking University) · Xuanchen Ren () · Shuming Ma () · Houfeng Wang ()

MEC: Memory-efficient Convolution for Deep Neural Network
Minsik Cho (IBM Research) · Daniel Brand ()

Scaling Up Sparse Support Vector Machine by Simultaneous Feature and Sample Reduction
Weizhong Zhang (Zhejiang University & Tencent Inc) · Bin Hong (Zhejiang University) · Jieping Ye (University of Michigan) · Deng Cai (Zhejiang University) · Xiaofei He (Zhejiang University) · Jie Wang (University of Michigan)

Bayesian inference on random simple graphs with power law degree distributions
Juho Lee (POSTECH) · Creighton Heaukulani (Cambridge University) · Lancelot James (HKUST) · Seungjin Choi (POSTECH) · Zoubin Ghahramani (University of Cambridge)

Globally Optimal Gradient Descent for a ConvNet with Gaussian Inputs
Alon Brutzkus (Tel Aviv University) · Amir Globerson (Tel Aviv University)

Coupling Distributed and Symbolic Execution for Natural Language Queries
Lili Mou (Peking University) · Zhengdong Lu (DeeplyCurious) · Hang Li (Huawei) · Zhi Jin (Peking University)

Learning from Clinical Judgments: Semi-Markov-Modulated Marked Hawkes Processes for Risk Prognosis
Ahmed M. Alaa Ibrahim (UCLA) · Scott Hu (UCLA) · Mihaela van der Schaar (Oxford University)

Learning Discrete Representations via Information Maximizing Self-Augmented Training
Weihua Hu (The University of Tokyo) · Takeru Miyato (Preferred Networks, Inc., ATR) · Seiya Tokui (Preferred Networks/ The University of Tokyo) · Eiichi Matsumoto (Preferred Networks Inc.) · Masashi Sugiyama (RIKEN / The University of Tokyo)

Multiplicative Normalizing Flows for Variational Bayesian Neural Networks
Christos Louizos (University of Amsterdam) · Max Welling (University of Amsterdam)

Random Feature Expansions for Deep Gaussian Processes
Kurt Cutajar (EURECOM) · Edwin Bonilla (The University of New South Wales) · Pietro Michiardi () · Maurizio Filippone (Eurecom)

A Laplacian Framework for Option Discovery in Reinforcement Learning
Marlos C. Machado (University of Alberta) · Marc Bellemare (DeepMind) · Michael Bowling (University of Alberta)

Gradient Projection Iterative Sketch for Large-scale Constrained Least-squares
Junqi Tang (the University of Edinburgh) · Mohammad Golbabaee (the University of Edinburgh) · Mike Davies (the University of Edinburgh)

Innovation Pursuit: A New Approach to the Subspace Clustering Problem
Mostafa Rahmani (University of Central Florida) · George Atia (University of Central Florida)

A Distributional Perspective on Reinforcement Learning
Marc Bellemare (DeepMind) · Will Dabney (DeepMind) · Remi Munos (Google DeepMind)

Efficient Algorithms for Online Non-Convex Optimization
Elad Hazan (Princeton University) · Karan Singh (Princeton University) · Cyril Zhang (Princeton University)

The Price of Differential Privacy For Online Learning
Naman Agarwal (Princeton University) · Karan Singh (Princeton University)

On Context-Dependent Clustering of Bandits
Claudio Gentile (Universita dell'Insubria) · Shuai Li (University of Cambridge) · Purushottam Kar (Indian Institute of Technology Kanpur) · Alexandros Karatzoglou (Telefonica Research) · Giovanni Zappella (Amazon Dev Center Germany) · Evans Howard (University of Insubria)

Efficient Distributed Learning with Sparsity
Jialei Wang (University of Chicago) · Mladen Kolar (University of Chicago) · Nati Srebro (Toyota Technological Institute at Chicago) · Tong Zhang ()

A Simulated Annealing Based Inexact Oracle for Wasserstein Loss Minimization
Jianbo Ye (Penn State University) · James Wang (Penn State University) · Jia Li (Penn State University)

End-to-End Differentiable Adversarial Imitation Learning
Nir Baram (Technion) · Oron Anschel (Technion) · Itai Caspi (Technion) · Shie Mannor (Technion)

Dueling Bandits with Weak Regret
Bangrui Chen (Cornell University) · Peter Frazier (Cornell University)

Consistent k-Clustering
Silvio Lattanzi () · Sergei Vassilvitskii (Google)

(Even More) Efficient Reinforcement Learning via Posterior Sampling
Ian Osband (Deepmind) · Benjamin Van Roy (Stanford)

Statistical Inference for Incomplete Ranking Data: The Case of Rank-Dependent Coarsening
Mohsen Ahmadi Fahandar (University of Paderborn) · Eyke Hullermeier () · Ines Couso (University of Oviedo)

Co-clustering through Optimal Transport
Charlotte Laclau (LIG) · Ievgen Redko () · Basarab Matei () · Younès Bennani () · Vincent Brault (Univ. Grenoble Alpes)

Just Sort It! A Simple and Effective Approach to Active Preference Learning
Lucas Maystre (EPFL) · Matthias Grossglauser (EPFL)

Depth-Width Tradeoffs in Approximating Natural Functions With Neural Networks
Itay Safran (Weizmann Institute of Science) · Ohad Shamir (Weizmann Institute of Science)

Natasha: Faster Non-Convex Stochastic Optimization Via Strongly Non-Convex Parameter
Zeyuan Allen-Zhu (Institute for Advanced Study)

Nyström Method with Kernel K-Means++ Samples as Landmarks
Dino Oglic (University of Bonn) · Thomas Gaertner (The University of Nottingham)

Multi-fidelity Bayesian Optimisation with Continuous Approximations
kirthevasan kandasamy (CMU) · Gautam Dasarathy (Rice University) · Barnabás Póczos (CMU) · Jeff Schneider ()

Graph-based Isometry Invariant Representation Learning
Renata Khasanova (Ecole Polytechnique Federale de Lausanne (EPFL)) · Pascal Frossard (EPFL)

Improved multitask learning through synaptic intelligence
Friedemann Zenke (Stanford) · Ben Poole (Stanford University) · Surya Ganguli (Stanford)

Strongly-Typed Agents are Guaranteed to Interact Safely
David Balduzzi (Victoria University Wellington)

Neural Taylor Approximations: Convergence and Exploration in Rectifier Networks
David Balduzzi (Victoria University Wellington) · Brian McWilliams (Disney Research) · Tony Butler-Yeoman (Victoria University of Wellington)

The Shattered Gradients Problem: If resnets are the answer, then what is the question?
David Balduzzi (Victoria University Wellington) · Marcus Frean (Victoria University Wellington) · Wan-Duo Ma (Victoria University) · Brian McWilliams (Disney Research) · Lennox Leary (VUW) · J.P. Lewis (Frostbite Labs and Victoria University)

On Mixed Memberships and Symmetric Nonnegative Matrix Factorizations
Xueyu Mao (University of Texas at Austin) · Purnamrita Sarkar (UT Austin) · Deepayan Chakrabart ()

Semi-Supervised Classification Based on Classification from Positive and Unlabeled Data
Tomoya Sakai (The University of Tokyo) · Marthinus du Plessis (Indeed) · Gang Niu (University of Tokyo) · Masashi Sugiyama (RIKEN / The University of Tokyo)

Rule-Enhanced Penalized Regression by Column Generation using Rectangular Maximum Agreement
Jonathan Eckstein (Rutgers University) · Noam Goldberg (Bar-Ilan University) · Ai Kagawa (Rutgers Univeristy)

SARAH: A Novel Method for Machine Learning Problems Using Stochastic Recursive Gradient
Lam Nguyen (Lehigh University) · Jie Liu (Lehigh University) · Katya Scheinberg (Lehigh University) · Martin Takac (Lehigh)

PixelCNN models with Auxiliary Variables for Natural Image Modeling
Alexander Kolesnikov (IST Austria) · Christoph Lampert (IST Austria)

Sharp Minima Can Generalize For Deep Nets
Laurent Dinh (U. Montreal) · Razvan Pascanu (DeepMind) · Samy Bengio (Google Brain) · Yoshua Bengio (U. Montreal)

Evaluating the Variance of Likelihood-Ratio Gradient Estimators
Seiya Tokui (Preferred Networks/ The University of Tokyo) · Issei Sato (University of Tokyo)

Near-Optimal Design of Experiments via Regret Minimization
Zeyuan Allen-Zhu (Institute for Advanced Study) · Yuanzhi Li (Princeton University) · Aarti Singh () · Yining Wang (CMU)

Contextual Decision Processes with low Bellman rank are PAC-Learnable
Nan Jiang (Microsoft Research) · Akshay Krishnamurthy (UMass) · Alekh Agarwal (Microsoft Research) · John Langford (Microsoft Research) · Robert Schapire (Microsoft Research)

Differentially Private Ordinary Least Squares
Or Sheffet (University of Alberta)

Differentially Private Learning of Graphical Models using CGMs
Garrett Bernstein (University of Massachusetts Amherst) · Ryan McKenna () · Tao Sun () · Michael Hay () · Gerome Miklau () · Daniel Sheldon (UMass Amherst)

Leveraging Union of Subspace Structure to Improve Constrained Clustering
Laura Balzano (University of Michigan) · John Lipor (University of Michigan)

Learning Important Features Through Propagating Activation Differences
(None) · Peyton Greenside (Stanford University) · Anshul Kundaje (Stanford University)

Probabilistic Path Hamiltonian Monte Carlo
Vu Dinh (Fred Hutchinson Cancer Center) · Arman Bilge (University of Washington) · Cheng Zhang (Fred Hutchinson Cancer Center) · Frederick Matsen (Fred Hutchinson Cancer Center)

Asymmetric Tri-training for Unsupervised Domain Adaptation
Saito Kuniaki (The University of Tokyo) · Yoshitaka Ushiku (The University of Tokyo) · Tatsuya Harada ()

Logarithmic Time One-Against-Some
Hal Daumé (University of Maryland) · NIKOS KARAMPATZIAKIS (Microsoft) · John Langford (Microsoft Research) · Paul Mineiro (Microsoft)

Optimal and Adaptive Off-policy Evaluation in Contextual Bandits
Yu-Xiang Wang (CMU) · Alekh Agarwal (Microsoft Research) · Miroslav Dudik (Microsoft Research)

Identifying Best Interventions through Online Importance Sampling
Rajat Sen (University of Texas at Austin) · Karthikeyan Shanmugam (IBM Thomas J. Watson Research Center) · Sanjay Shakkottai (University of Texas at Austin) · Alexandros Dimakis (UT Austin)

Analogical Inference for Multi-relational Embeddings
Hanxiao Liu (Carnegie Mellon University) · Yiming Yang (Carnegie Mellon University) · Yuexin Wu (Carnegie Mellon University)

Coordinated Multi-Agent Imitation Learning
Hoang Le (Caltech) · Yisong Yue (Caltech) · Peter Carr (Disney Research)

Fast Bayesian Intensity Estimation for the Permanental Process
Christian Walder (Data61) · Adrian Bishop (Data61/ANU/UTS)

Sequence to Better Sequence: Continuous Revision of Combinatorial Structures
Jonas Mueller (MIT) · David Gifford (MIT) · Tommi Jaakkola (MIT)

A Universal Variance Reduction-Based Framework for Nonconvex Low-Rank Matrix Recovery
Lingxiao Wang (University of Virginia) · Xiao Zhang (University of Virginia) · Quanquan Gu (University of Virginia)

Zero-Inflated Exponential Family Embeddings
Liping Liu (Columbia University) · David Blei (Columbia University)

Clustering High Dimensional Dynamic Data Streams
Lin Yang (Johns Hopkins) · Harry Lang (Johns Hopkins University) · Christian Sohler (TU Dortmund) · Vladimir Braverman (Johns Hopkins University) · Gereon Frahling (Linguee GmbH)

Optimal Densification for Fast and Accurate Minwise Hashing
Anshumali Shrivastava ()

Safety-Aware Algorithms for Adversarial Contextual Bandit
Wen Sun (Carnegie Mellon University) · Debadeepta Dey (Microsoft) · Ashish Kapoor (Microsoft Research)

Asynchronous Distributed Variational Gaussian Processes
Hao Peng (Purdue University) · Shandian Zhe (Purdue University) · Xiao Zhang (Purdue University) · Yuan Qi (Ant Financial)

Max-value Entropy Search for Efficient Bayesian Optimization
Zi Wang (MIT) · Stefanie Jegelka (MIT)

Tensor Balancing on Statistical Manifold
Mahito Sugiyama (National Institute of Informatics) · Hiroyuki Nakahara (RIKEN Brain Science Institute) · Koji Tsuda (University of Tokyo)

Adaptive Consensus ADMM for Distributed Optimization
Zheng Xu (University of Maryland) · Gavin Taylor (US Naval Academy) · Hao Li (University of Maryland at College Park) · Mario Figueiredo (Instituto Superior Tecnico) · Xiaoming Yuan () · Tom Goldstein ()

Coherent probabilistic forecasts for hierarchical time series
Souhaib Ben Taieb (Monash University) · James Taylor (University of Oxford) · Rob Hyndman (Monash University)

Large-Scale Evolution of Image Classifiers
Esteban Real (Google Inc.) · Sherry Moore (Google Inc.) · Andrew Selle (Google Inc.) · Saurabh Saxena (Google Inc.) · Yutaka Suematsu (Google Inc.) · Quoc Le (Google Brain) · Alex Kurakin (Google Inc.)

Provable Alternating Gradient Descent for Non-negative Matrix Factorization with Strong Correlations
Yuanzhi Li (Princeton University) · Yingyu Liang (Princeton University)

Convex Relaxation without Lifting
Tom Goldstein () · Christoph Studer (Cornell University)

Differentially Private Submodular Maximization: Data Summarization in Disguise
Marko Mitrovic (Yale University) · Mark Bun (Princeton University) · Andreas Krause (ETH Zurich) · Amin Karbasi (Yale)

Faster Greedy MAP Inference for Determinantal Point Processes
Insu Han (Korea Advanced Institute of Science and Technology) · Jinwoo Shin (KAIST) · Kyoungsoo Park (KAIST) · Prabhanjan Kambadur (Bloomberg)

Diffusion Independent Semi-Bandit Influence Maximization
Sharan Vaswani (UBC) · Branislav Kveton (Adobe Research) · Zheng Wen (Adobe Research) · Mohammad Ghavamzadeh (Adobe) · Laks Lakshmanan (University of British Columbia) · Mark Schmidt (University of British Columbia)

How to Escape Saddle Points Efficiently
Chi Jin (UC Berkeley) · Rong Ge (Duke University) · Praneeth Netrapalli (Microsoft Research) · Sham M. Kakade (University of Washington) · Michael Jordan ()

Learning to Generate Long-term Future via Hierarchical Prediction
Ruben Villegas (University of Michigan) · Jimei Yang () · Xunyu Lin () · Yuliang Zou (University of Michigan) · Sungryull Sohn (University of Michigan) · Honglak Lee (Google / U. Michigan)

Deciding How to Decide: Dynamic Routing in Artificial Neural Networks
Mason McGill (Caltech)

Parallel Multiscale Autoregressive Density Estimation
Scott Reed (Google Deepmind) · Aäron van den Oord (Google) · Nal Kalchbrenner (DeepMind) · Sergio Gomez (Google) · Ziyu Wang (Deep Mind) · Dan Belov (Google) · Nando de Freitas (DeepMind)

Graphical Models for Ordinal Data: A Tale of Two Approaches
ARUN SAI SUGGALA (Carnegie Mellon University) · Eunho Yang (Korea Advanced Institute of Technology) · Pradeep Ravikumar (Carnegie Mellon University)

Online Learning to Rank in Stochastic Click Models
Mohammad Ghavamzadeh (Adobe) · Branislav Kveton (Adobe Research) · Csaba Szepesvari (University of Alberta) · Tomas Tunys (Czech Technical University) · Zheng Wen (Adobe Research) · Masrour Zoghi (Independent Researcher)

Deep Voice: Real-time Neural Text-to-Speech
Andrew Gibiansky (Baidu Research Silicon Valley AI Lab) · Mike Chrzanowski (Baidu Research) · Mohammad Shoeybi (Baidu Research) · Shubho Sengupta (Baidu Research) · Gregory Diamos (Baidu Research) · Sercan Arik (Baidu Research) · Jonathan Raiman (Baidu Research) · John Miller (Baidu Research) · Xian Li (Baidu) · Yongguo Kang (Baidu)

Sparse + Group-Sparse Dirty Models: Statistical Guarantees without Unreasonable Conditions and a Case for Non-Convexity
Eunho Yang (Korea Advanced Institute of Technology) · Aurelie Lozano (IBM)

Stochastic Variance Reduction Methods for Policy Evaluation
Simon Du (Carnegie Mellon University) · Jianshu Chen (Microsoft Research) · Lihong Li (Microsoft Research) · Lin Xiao (Microsoft Research) · Dengyong Zhou (Microsoft Research)

An Infinite Hidden Markov Model With Similarity-Biased Transitions
Colin Dawson (Oberlin College) · Chaofan Huang (Oberlin College) · Clayton Morrison (University of Arizona)

Algorithmic stability and hypothesis complexity
Tongliang Liu (University of Sydney) · Gábor Lugosi (Universitat Pompeu Fabra) · Gergely Neu () · Dacheng Tao ()

Tensor Belief Propagation
Andrew Wrigley (Australian National University) · Wee Lee (National University of Singapore) · Nan Ye (Queensland University of Technology)

Schema Networks
Ken Kansky (Vicarious Systems FPC, Inc.) · David Mely (Vicarious Systems) · Mohamed Eldawy (Vicarious Systems) · Thomas Silver (Vicarious) · Miguel Lazaro-Gredilla (Vicarious) · Xinghua Lou (Vicarious Systems) · Nimrod Dorfman (Vicarious Systems) · Dileep George (Vicarious) · Scott Phoenix (Vicarious Systems)

Dance Dance Convolution
Christopher Donahue (University of California, San Diego) · Zachary Lipton (UCSD) · Julian Mcauley (UCSD)

Provable Optimal Algorithms for Generalized Linear Contextual Bandits
Lihong Li (Microsoft Research) · Yu Lu (Yale University) · Dengyong Zhou (Microsoft Research)

Geometry of Neural Network Loss Surfaces via Random Matrix Theory
Jeffrey Pennington (Google Brain) · Yasaman Bahri ()

Recurrent Highway Networks
Julian Zilly (ETH Zurich) · Rupesh Srivastava (IDSIA (University of Lugano)) · Jan Koutnik (NNAISSENSE) · Jürgen Schmidhuber (Swiss AI Lab)

Prediction and Control with Temporal Segment Models
Nikhil Mishra (UC Berkeley) · Pieter Abbeel (OpenAI / Berkeley) · Igor Mordatch (OpenAI)

Learning Continuous Semantic Representations of Symbolic Expressions
Miltiadis Allamanis (Microsoft Research) · pankajan Chanthirasegaran () · Pushmeet Kohli (Microsoft Research) · Charles Sutton (University of Edinburgh)

Training Models with End-to-End Low Precision: The Cans, the Cannots, and a Little Bit of Deep Learning
Hantian Zhang (ETH Zurich) · Jerry Li (MIT) · Kaan Kara (ETH Zurich) · Dan Alistarh (ETH Zurich) · Ji Liu () · Ce Zhang (ETH Zurich)

Warped Convolutions: Efficient Invariance to Spatial Transformations
Joao Henriques (University of Oxford) · Andrea Vedaldi (University of Oxford)

RobustFill: Neural Program Learning under Noisy I/O
Jacob Devlin (Microsoft Research) · Jonathan Uesato (MIT) · Surya Bhupatiraju (MIT) · Rishabh Singh (Microsoft Research) · Abdelrahman Mohammad (Microsoft) · Pushmeet Kohli (Microsoft Research)

Dictionary Learning Based on Sparse Distribution Tomography
Pedram Pad () · Farnood Salehi (EPFL) · Elisa Celis () · Patrick Thiran (EPFL) · Michael Unser ()

On the Iteration Complexity of Support Recovery via Hard Thresholding Pursuit
Jie Shen (Rutgers University) · Ping Li (Rugters University)

Learning Texture Manifolds with the Periodic Spatial GAN
Nikolay Jetchev (Zalando Research) · Urs M Bergmann (Zalando Research) · Roland Vollgraf (Zalando Research)

Decoupled Neural Interfaces using Synthetic Gradients
Max Jaderberg (DeepMind) · Wojciech Czarnecki (DeepMind) · Simon Osindero (DeepMind) · Oriol Vinyals (DeepMind) · Alex Graves (Google DeepMind) · David Silver (Google DeepMind) · Koray Kavukcuoglu (DeepMind)

Bayesian Optimization with Tree-structured Dependencies
Rodolphe Jenatton (Amazon) · Cedric Archambeau (Amazon) · Javier González (Amazon) · Matthias Seeger (Amazon.com)

Robust Budget Allocation via Continuous Submodular Functions
Matthew Staib (MIT) · Stefanie Jegelka (MIT)

Adapting kernel representations online using submodular maximization
Yangchen Pan (Indiana University) · Matthew Schlegel (Indiana University) · Jiecao Chen (Indiana University Bloomington) · Martha White (Indiana University)

Minimizing Trust Leaks for Robust Sybil Detection
Janos Höner (TU Berlin) · Alexander Bauer (TU Berlin) · Klaus-robert Mueller () · Shinichi Nakajima (TU Berlin) · Nico Goernitz (TU Berlin)

Efficient Orthogonal Parametrisation of Recurrent Neural Networks Using Householder Reflections
zakaria mhammedi (The University of Melbourne) · Andrew Hellicar (CSIRO) · James Bailey (The University of Melbourne) · Ashfaqur Rahman (CSIRO)

Adversarial Variational Bayes: Unifying Variational Autoencoders and Generative Adversarial Networks

Lars Mescheder (MPI Tübingen) · Sebastian Nowozin (Microsoft Research) · Andreas Geiger (MPI Tübingen)

A new formulation for deep ordinal classification
Christopher Beckham (MILA) · Christopher Pal (MILA)

Uncovering Causality from Multivariate Hawkes Integrated Cumulants
Massil Achab (Ecole Polytechnique) · Emmanuel Bacry (Ecole Polytechnique) · Stéphane Gaïffas (CMAP CNRS UMR 7641) · Iacopo Mastromatteo (Capital Fund Management) · Jean-François Muzy (Université de Corse)

Robust Submodular Maximization: A Non-Uniform Partitioning Approach
Ilija Bogunovic (EPFL) · Slobodan Mitrovic (EPFL) · Jonathan Scarlett (EPFL) · Volkan Cevher (EPFL)

A Simple Multi-Class Boosting Framework with Theoretical Guarantees and Empirical Proficiency
Ron Appel (Caltech.edu) · Pietro Perona (caltech.edu)

Boosted Fitted Q-Iteration
Marcello Restelli (Politecnico di Milano) · Matteo Pirotta (INRIA) · Carlo D'Eramo (Politecnico di Milano) · Samuele Tosatto (Politecnico di Milano)

Multi-objective Bandits: Optimizing the Generalized Gini Index
Paul Weng () · Balazs Szorenyi () · Shie Mannor (Technion) · Robert Busa-Fekete (Yahoo! Research)

Understanding Black-box Predictions via Influence Functions
Pang Koh (Stanford University) · Percy Liang (Stanford University)

Source-Target Similarity Modelings for Multi-Source Transfer Gaussian Process Regression
PENGFEI WEI (Nanyang Technological University, Singapore) · Ramon Sagarna () · Yiping Ke () · CHI GOH () · yEW ONG ()

Zonotope hit-and-run for efficient sampling from projection DPPs
Guillaume Gautier (INRIA Lille) · Rémi Bardenet (CNRS and Univ. Lille) · Michal Valko (Inria Lille - Nord Europe)

Identify the Nash Equilibrium in Static Games with Random Payoffs
Yichi Zhou (Tsinghua University) · Jialian Li (Tsinghua University) · Jun Zhu (Tsinghua University)

AdaNet: Adaptive Structural Learning of Artificial Neural Networks
Corinna Cortes (Google Research) · Xavi Gonzalvo () · Vitaly Kuznetsov (Google) · Mehryar Mohri (Courant Institute and Google Research) · Scott Yang (Courant Institute)

ProtoNN: Compressed and Accurate kNN for Resource-scarce Devices
Chirag Gupta (Microsoft Research, India) · ARUN SUGGALA (Carnegie Mellon University) · Ankit Goyal (University of Michigan) · Saurabh Goyal (IIT Delhi) · Ashish

Kumar (Microsoft Research) · Bhargavi Paranjape (Microsoft Research) · Harsha Simhadri (Microsoft Research) · Raghavendra Udupa (Microsoft Research) · (None) · Prateek Jain (Microsoft Research)

The Statistical Recurrent Unit
Junier Oliva (CMU) · Barnabás Póczos (CMU) · Jeff Schneider ()

Optimal algorithms for smooth and strongly convex distributed optimization in networks
Kevin Scaman (MSR-INRIA Joint Center) · Yin Lee (Microsoft Research) · Francis Bach (INRIA) · Sebastien Bubeck (Microsoft Research) · Laurent Massoulié (MSR-INRIA Joint Center)

Equivariance Through Parameter-Sharing
Mohsen Ravanbakhsh (Carnegie Mellon University) · Jeff Schneider (CMU) · Barnabás Póczos (CMU)

Learning to learn without gradient descent by gradient descent
Yutian Chen (Deep Mind) · Matthew Hoffman (DeepMind) · Sergio Gomez (Google) · Misha Denil (University of Oxford) · Timothy Lillicrap (Google DeepMind) · Matthew Botvinick (DeepMind) · Nando de Freitas (DeepMind)

Local-to-Global Bayesian Network Structure Learning
Tian Gao (IBM) · Kshitij Fadnis (IBM) · Murray Campbell (IBM)

Distributed Batch Gaussian Process Optimization
Erik A. Daxberger (Ludwig-Maximilians-Universität München) · Bryan Kian Hsiang Low (National University of Singapore)

Multi-task Learning with Labeled and Unlabeled Tasks
Anastasia Pentina (IST Austria) · Christoph Lampert (IST Austria)

SPLICE: Fully Tractable Hierarchical Extension of ICA with Pooling
Jun-ichiro Hirayama (RIKEN AIP / ATR) · Aapo Hyvärinen (UCL) · Motoaki Kawanabe (ATR)

A birth-death process for feature allocation
Konstantina Palla (Oxford) · David Knowles (Stanford) · Zoubin Ghahramani (University of Cambridge)

Confident Multiple Choice Learning
Kimin Lee (KAIST) · Jinwoo Shin (KAIST) · Changho Hwang (KAIST) · KyoungSoo Park (KAIST)

Failures of Gradient-Based Deep Learning
Shaked Shammah (Hebrew University Jerusalem Israel) · Shai Shalev-Shwartz () · Ohad Shamir (Weizmann Institute of Science)

On the Sampling Problem for Kernel Quadrature
Francois-Xavier Briol (University of Warwick) · Chris J Oates (Newcastle University) · Jon Cockayne (University of Warwick) · Mark Girolami (Imperial College London)

Resource-efficient Machine Learning in 2 KB RAM for the Internet of Things
Ashish Kumar (Microsoft Research) · Saurabh Goyal (IIT Delhi) · (None)

Fairness in Reinforcement Learning
Shahin Jabbari (University of Pennsylvania) · Matthew Joseph (University of Pennsylvania) · Michael Kearns (University of Pennsylvania) · Jamie Morgenstern (University of Pennsylvania) · Aaron Roth (University of Pennsylvania)

Deletion-Robust Submodular Maximization: Data Summarization with "the Right to be Forgotten"
Baharan Mirzasoleiman (ETH Zurich) · Amin Karbasi (Yale) · Andreas Krause (ETH Zurich)

Clustering by Sum of Norms: Stochastic Incremental Algorithm, Convergence and Cluster Recovery
Ashkan Panahi (NC state university) · devdatt Dubhashi (Chalmers University) · Fredrik D Johansson (MIT) · Chiranjib Bhattacharya ()

Projection-Free Distributed Online Learning in Networks
Wenpeng Zhang (Tsinghua University) · Peilin Zhao () · Wei Liu (Tencent AI Lab) · Steven Hoi (Singapore Management University) · wenwu zhu () · Tong Zhang ()

Automated Curriculum Learning for Neural Networks
Alex Graves (Google DeepMind) · Marc Bellemare (DeepMind) · Jacob Menick (DeepMind) · Remi Munos (Google DeepMind) · Koray Kavukcuoglu (DeepMind)

Meta Networks
Tsendsuren Munkhdalai (University of Massachusetts) · Hong Yu (University of Massachusetts)

Deep Latent Dirichlet Allocation with Topic-Layer-Adaptive Stochastic Gradient Riemannian MCMC
Yulai Cong (Xidian University) · Bo Chen (National Lab of Radar Signal Processing, School of Electronic Engineering, Xidian University) · Hongwei Liu (Xidian University) · Mingyuan Zhou (University of Texas at Austin)

Forward and Reverse Gradient-Based Hyperparameter Optimization
Luca Franceschi (IIT and UCL) · Michele Donini (IIT) · Paolo Frasconi (University of Florence) · Massimiliano Pontil (University College London)

McGan: Mean and Covariance Feature Matching GAN
Youssef Mroueh (IBM) · Tom Sercu (IBM) · Vaibhava Goel (IBM)

Learning to Discover Sparse Graphical Models
Eugene Belilovsky (CentraleSupelec) · Kyle Kastner () · Gael Varoquaux () · Matthew Blaschko (KU Leuven)

The Predictron: End-To-End Learning and Planning
David Silver (Google DeepMind) · Hado van Hasselt (DeepMind) · Matteo Hessel (Deep Mind) · Tom Schaul (DeepMind) · Arthur Guez (Google DeepMind) · Tim Harley

() · Gabriel Dulac-Arnold (Google DeepMind) · David Reichert (DeepMind) · Neil Rabinowitz (DeepMind) · Andre Barreto (Google DeepMind) · Thomas Degris ()

A Generative Framework for Multi-label Learning with Missing Labels
Vikas Jain (Indian Institute of Technology Kanpur) · Nirbhay Modhe () · Piyush Rai (IIT Kanpur)

Deeply AggreVaTeD: Differentiable Imitation Learning for Sequential Prediction
Wen Sun (Carnegie Mellon University) · Arun Venkatraman (Carnegie Mellon University) · Geoff Gordon (Carnegie Mellon University) · Byron Boots (Georgia Tech) · J. Bagnell (Carnegie Mellon University)

Algorithms for ℓ_p Low-Rank Approximation
Flavio Chierichetti (Sapienza University of Rome) · Sreenivas Gollapudi () · Ravi Kumar (Google) · Silvio Lattanzi () · Rina Panigrahy () · David Woodruff ()

DARLA: Improving Zero-Shot Transfer in Reinforcement Learning
Irina Higgins (DeepMind) · Arka Pal (DeepMind) · Andrei Rusu (DeepMind) · Loic Matthey (DeepMind) · Christopher Burgess (DeepMind) · Alexander Pritzel (Deepmind) · Matthew Botvinick (DeepMind) · Charles Blundell (DeepMind) · Alexander Lerchner (DeepMind)

Hierarchical Latent Feature Models for Relational Data with Side Information
Changwei Hu (Duke University) · Piyush Rai (IIT Kanpur) · Lawrence Carin (Duke)

Multilabel Classification with Group Testing and Codes
Shashanka Ubaru (University of Minnesota) · Arya Mazumdar (University of Massachusetts at Amherst)

Distributed Mean Estimation with Limited Communication
Ananda Suresh (Google Research) · Felix Yu (Google Research) · Sanjiv Kumar (Google Research, NY) · Brendan McMahan (Google)

Approximate Newton Methods and Their Local Convergence
Haishan Ye (Shanghai Jiao Tong University) · Luo Luo (Shanghai Jiao Tong University) · Zhihua Zhang ()

Bayesian Boolean Matrix Factorisation
Tammo Rukat (University of Oxford) · Christopher Holmes (University of Oxford) · Michalis Titsias (Athens University of Economics and Business) · Christopher Yau (University of Oxford)

Global optimization of Lipschitz functions
Cédric Malherbe (ENS Paris-Saclay) · Nicolas Vayatis (ENS Cachan)

Robust Gaussian Graphical Model Estimation with Arbitrary Corruption
Lingxiao Wang (University of Virginia) · Quanquan Gu (University of Virginia)

Understanding Synthetic Gradients and Decoupled Neural Interfaces
Wojciech Czarnecki (DeepMind) · Grzegorz Świrszcz (DeepMind) · Max Jaderberg

(DeepMind) · Simon Osindero (DeepMind) · Oriol Vinyals (DeepMind) · Koray Kavukcuoglu (DeepMind)

Video Pixel Networks
Nal Kalchbrenner (DeepMind) · Karen Simonyan (DeepMind) · Aäron van den Oord (Google) · Ivo Danihelka (Google DeepMind) · Oriol Vinyals (DeepMind) · Alex Graves (Google DeepMind) · Koray Kavukcuoglu (DeepMind)

Learning Determinantal Point Processes with Moments and Cycles
John C Urschel (Massachusetts Institute of Technology) · Victor Brunel () · Ankur Moitra () · Philippe Rigollet (MIT)

Frame-based Data Factorizations
Sebastian Mair (Leuphana University Lüneburg) · Ahcène Boubekki (Leuphana University) · Ulf Brefeld (Leuphana University)

Approximate Steepest Coordinate Descent
Sebastian Stich (EPFL) · Anant Raj (Max-Planck Institute for Intelligent Systems) · Martin Jaggi (EPFL)

The loss surface of deep and wide neural networks
Quynh Nguyen (Saarland University) · Matthias Hein (Saarland University)

Hierarchy Through Composition with Multitask LMDPs
Adam Earle (University of the Witwatersran) · Andrew Saxe (Harvard University) · Benjamin Rosman (Council for Scientific and Industrial Research (CSIR))

Strong NP-Hardness for Sparse Optimization with Concave Penalty Functions
(None) · Mengdi Wang (Princeton University) · Dongdong Ge () · Zizhuo Wang (University of Minnesota) · Yinyu Ye ()

Pain-Free Random Differential Privacy with Sensitivity Sampling
Benjamin Rubinstein (Univ of Melbourne) · Francesco Aldà (Ruhr-Universität Bochum)

Improving Viterbi is Hard: Better Runtimes Imply Faster Clique Algorithms
Arturs Backurs (MIT) · Christos Tzamos (MIT)

Exact MAP Inference by Avoiding Fractional Vertices
Erik Lindgren (University of Texas at Austin) · Alexandros Dimakis (UT Austin) · Adam Klivans (University of Texas at Austin)

Attentive Recurrent Comparators
Pranav Shyam Manjunatha (R. V. College of Engineering) · Shubham Gupta () · Ambedkar Dukkipati (Indian Institute of Science)

DeepBach: A Steerable Model for Bach Chorales Generation
Gaëtan HADJERES (LIP6 / SONY CSL) · François Pachet (Sony CSL / UPMC) · Frank Nielsen (Sony CSL)

Survival HMM: An Interpretable, Event-time Prediction Model for mHealth
Walter Dempsey (University of Michigan) · Alexander Moreno (Georgia Institute of Technology) · Jim Rehg (Georgia Tech) · Susan Murphy (University of Michigan)

Improving Stochastic Policy Gradients in Continuous Control with Deep Reinforcement Learning using the Beta Distribution
Po-Wei Chou (Carnegie Mellon University) · Daniel Maturana (Carnegie Mellon University) · Sebastian Scherer (Carnegie Mellon University)

Multichannel End-to-end Speech Recognition
Tsubasa Ochiai (Doshisha University) · Shinji Watanabe (MITSUBISHI ELECTRIC RESEARCH LABORATORIES) · Takaaki Hori (MITSUBISHI ELECTRIC RESEARCH LABORATORIES) · John Hershey (MITSUBISHI ELECTRIC RESEARCH LABORATORIES)

Scalable Bayesian Rule Lists
Hongyu Yang (Massachusetts Institute of Technology) · Cynthia Rudin (Duke University) · Margo Seltzer (Harvard University)

Hyperplane Clustering Via Dual Principal Component Pursuit
Manolis Tsakiris (Johns Hopkins University) · Rene Vidal (Johns Hopkins University)

High-dimensional Non-Gaussian Single Index Models via Thresholded Score Function Estimation
Zhuoran Yang (Princeton University) · Krishnakumar Balasubramanian (Princeton) · Han Liu (Princeton University)

Towards K-means-friendly Spaces: Simultaneous Deep Learning and Clustering
Bo Yang (University of Minnesota) · Xiao Fu (University of Minnesota) · Nicholas Sidiropoulos (University of Minnesota) · Mingyi Hong (Iowa State University)

Batched High-dimensional Bayesian Optimization via Structural Kernel Learning
Zi Wang (MIT) · Chengtao Li () · Stefanie Jegelka (MIT) · Pushmeet Kohli (Microsoft Research)

On orthogonality and learning RNNs with long term dependencies
Eugene Vorontsov (Ecole Polytechnique de Montrea) · Chiheb Trabelsi (Ecole Polytechnique de Montreal) · Christopher Pal (École Polytechnique de Montréal) · Samuel Kadoury (Ecole Polytechnique de Montreal)

High-Dimensional Structured Quantile Regression
Vidyashankar Sivakumar (University of Minnesota) · Arindam Banerjee (University of Minnesota)

Gram-CTC: Automatic Unit Selection and Target Decomposition for Sequence Labelling
Hairong Liu (Baidu Silicon Valley AI Lab) · Zhenyao Zhu (Baidu Silicon Valley AI Lab) · Xiangang Li (Baidu AI Lab) · Sanjeev Satheesh (Baidu SVAIL)

Analytical Guarantees on Numerical Precision of Deep Neural Networks
Charbel Sakr (University of Illinois at Urbana-Champaign) · Yongjune Kim (UIUC) · Naresh Shanbhag (University of Illinois)

Meritocratic Fairness for Cross-Population Selection
Steven Wu (UPenn) · Michael Kearns (University of Pennsylvania) · Aaron Roth (University of Pennsylvania)

Neural Episodic Control
Alexander Pritzel (Deepmind) · Benigno Uria (Deepmind) · Srinivasan Sriram (DeepMind) · Adrià Puigdomenech Badia (Deepmind) · Oriol Vinyals (DeepMind) · Daan Wierstra (Google DeepMind) · Charles Blundell (DeepMind)

Latent Intention Dialogue Models
Tsung-Hsien Wen (University of Cambridge) · Yishu Miao (University of Oxford) · Phil Blunsom (Oxford University and DeepMind) · Steve Young (University of Cambridge)

Cost-Optimal Learning of Causal Graphs
Murat Kocaoglu (University of Texas at Austin) · Alexandros Dimakis (UT Austin) · Sriram Vishwanath ()

Local Bayesian Optimization of Motor Skills
Riad Akrour (TU Darmstadt) · Dmitry Sorokin () · Jan Peters (TU Darmstadt) · Gerhard Neumann ()

Prox-PDA: The Proximal Primal-Dual Algorithm for Fast Distributed Nonconvex Optimization and Learning Over Networks
Mingyi Hong (Iowa State University) · Ming-Min Zhao (Zhejiang University) · Davood Hajinezhad (Iowa State University)

Learning in POMDPs with Monte Carlo Tree Search
Sammie Katt (Northeastern University) · Frans Oliehoek (University of Liverpool) · Chris Amato (Northeastern University)

A Unified View of Multi-Label Performance Measures
Xi-Zhu Wu (Nanjing University) · Zhi-Hua Zhou (Nanjing University)

Recovery Guarantees for One-hidden-layer Neural Networks
Kai Zhong (University of Texas at Austin) · Zhao Song (UT-Austin) · Prateek Jain (Microsoft Research) · Peter Bartlett (UC Berkeley) · Inderjit Dhillon (UT Austin & Amazon)

From Patches to Images: A Nonparametric Generative Model
Geng Ji (Brown University) · Michael Hughes (Harvard University) · Erik Sudderth (University of California)

Robust Adversarial Reinforcement Learning
Lerrel Pinto (Carnegie Mellon University) · James Davidson (Google Brain) · Rahul Sukthankar (Google Research) · Abhinav Gupta (Carnegie Mellon University)

Learning Infinite Layer Networks without the Kernel Trick
Roi Livni (Princeton) · Daniel Carmon (Tel-Aviv University) · Amir Globerson (Tel Aviv University)

Differentially Private Clustering in High-Dimensional Euclidean Spaces
Nina Balcan (Carnegie Mellon University) · Travis Dick (CMU) · Yingyu Liang (Princeton University) · Wenlong Mou (Peking University) · Hongyang Zhang (Carnegie Mellon University)

Accurate and Timely Real-time Prediction of Sepsis Using an End-to-end Multitask Gaussian Process RNN Classifier
Joseph Futoma (Duke University) · Sanjay Hariharan (Duke University) · Katherine Heller (Duke University)

Regularising Non-linear Models Using Feature Side-information
Maolaaisha Aminanmu (University of Geneva, HES) · Pablo Strasser (HES-UNIGE) · Alexandros Kalousis (HES-UNIGE)

Intelligible Language Modeling with Input Switched Affine Networks
Jakob Foerster (University of Oxford) · Justin Gilmer (Google Brain) · Jan Chorowski (Google Brain) · Jascha Sohl-Dickstein (Google Brain) · David Sussillo (Google Brain, Google Inc.)

Adaptive Feature Selection: Computationally Efficient Online Sparse Linear Regression under RIP
Satyen Kale (Google Research) · Zohar Karnin (yahoo) · Tengyuan Liang (UPenn) · David Pal (Yahoo)

Neural networks and rational functions
Matus Telgarsky (UIUC)

Efficient softmax approximation for GPUs
Edouard Grave () · Armand Joulin (Facebook) · Moustapha Cisse () · David Grangier (Facebook) · Herve Jegou (Facebook AI Research)

Dual Supervised Learning
Yingce Xia (USTC) · Tao Qin () · Wei Chen (Microsoft Research) · Jiang Bian (Microsoft) · Nenghai Yu (USTC) · Tie-Yan Liu (Microsoft)

StingyCD: Safely Avoiding Wasteful Updates in Coordinate Descent
Tyler Johnson (University of Washington) · Carlos Guestrin ()

Improving Gibbs Sampler Scan Quality with DoGS
Ioannis Mitliagkas (Stanford University) · Lester Mackey (Microsoft Research)

Stochastic Generative Hashing
Bo Dai (Georgia Tech) · Ruiqi Guo (Google Research) · Sanjiv Kumar (Google Research, NY) · Niao He (UIUC) · Le Song (Georgia Institute of Technology)

Parallel and Distributed Thompson Sampling for Large-scale Accelerated Exploration of Chemical Space

Jose Hernandez-Lobato (University of Cambridge) · James Requeima (University of Cambridge) · Edward Pyzer-Knapp (IBM) · alan Aspuru-Guzik ()

Stochastic Gradient Monomial Gamma Sampler
Yizhe Zhang (Duke university) · Changyou Chen (Duke) · zhe gan () · Ricardo Henao (Duke University) · Lawrence Carin (Duke)

Soft-DTW: a Differentiable Loss Function for Time-Series
Marco Cuturi (ENSAE / CREST) · Mathieu Blondel (NTT)

Tensor-Train Recurrent Neural Networks for Video Classification
Yinchong Yang (Ludwig-Maximilians-Universität München, Siemens AG) · Denis Krompass () · Volker Tresp (University of Munich)

Exact Inference for Integer Latent-Variable Models
Kevin Winner (University of Massachusetts, Amherst) · Debora Sujono () · Daniel Sheldon (UMass Amherst)

Nearly Optimal Robust Matrix Completion
Yeshwanth Cherapanamjeri (Microsoft Research) · Prateek Jain (Microsoft Research) · Kartik Gupta (Microsoft Research)

Adversarial Feature Matching for Text Generation
Yizhe Zhang (Duke university) · zhe gan () · Kai Fan () · Zhi Chen (Duke University) · Ricardo Henao (Duke University) · Lawrence Carin (Duke)

Minimax Regret Bounds for Reinforcement LEarning
Mohammad Gheshlaghi Azar (Deepmind) · Ian Osband (Google DeepMind) · Remi Munos (Google DeepMind)

Bayesian models of data streams with Hierarchical Power Priors
Andres Masegosa (University of Almeria) · Antonio Salmeron (University of Almeria) · Dario Ramos-Lopez (University of Almeria) · Helge Langseth (Norwegian University of Science and Technology) · Thomas Nielsen (Aalborg University)

Discovering Discrete Latent Topics with Neural Variational Inference
Yishu Miao (University of Oxford) · Edward Grefenstette (Deepmind) · Phil Blunsom (Oxford University and DeepMind)

Unified Optimization Landscape for Nonconvex Low Rank Problems
Rong Ge (Duke University) · Chi Jin (UC Berkeley) · Yi Zheng (Duke University)

Stabilising Experience Replay for Deep Multi-Agent Reinforcement Learning
Jakob Foerster (University of Oxford) · Nantas Nardelli (University of Oxford) · Gregory Farquhar (University of Oxford) · Phil Torr (Oxford) · Pushmeet Kohli (Microsoft Research) · Shimon Whiteson (University of Oxford)

On Kernelized Multi-armed Bandits
Sayak Ray Chowdhury (Indian Institute of Science) · Aditya Gopalan (Indian Institute of Science)

Learned Optimizers that Scale and Generalize
Olga Wichrowska (Google Brain) · Niru Maheswaranathan (Stanford) · Matthew Hoffman (DeepMind) · Sergio Gomez (Google) · Misha Denil (University of Oxford) · Nando de Freitas (DeepMind) · Jascha Sohl-Dickstein (Google Brain)

An Alternative Softmax Operator for Reinforcement Learning
Kavosh Asadi (Brown University) · Michael Littman (Brown University)

Simultaneous Learning of Trees and Representations for Extreme Classification and Density Estimation
Yacine Jernite (New York University) · Anna Choromanska (New York University) · David Sontag (Massachusetts Institute of Technology)

Identification and Model Testing in Linear Structural Equation Models using Auxiliary Variables
Bryant Chen (IBM Research) · Daniel Kumor (Purdue University) · Elias Bareinboim (Purdue)

Differentiable Programs with Neural Libraries
Alex Gaunt (Microsoft) · Marc Brockschmidt (Microsoft Research) · Nate Kushman (Microsoft Research) · Daniel Tarlow (Google Brain)

Active Heteroscedastic Regression
Kamalika Chaudhuri (University of California at San Diego) · Prateek Jain (Microsoft Research) · Nagarajan Natarajan (Microsoft Research)

Prediction under Uncertainty in Sparse Spectrum Gaussian Processes with Applications to Filtering and Control
Yunpeng Pan (Georgia Tech) · Xinyan Yan () · Evangelos Theodorou (Georgia Tech) · Byron Boots (Georgia Tech)

Consistency Analysis for Binary Classification Revisited
Wojciech Kotlowski () · Nagarajan Natarajan (Microsoft Research) · Krzysztof Dembczynski () · Oluwasanmi Koyejo (University of Illinois at Urbana-Champaign)

Multilevel Clustering via Wasserstein Means
Nhat Ho (University of Michigan) · XuanLong Nguyen (University of Michigan) · Mikhail Yurochkin (University of Michigan) · Hung Bui (Adobe Research) · Viet Huynh () · Dinh Phung (Deakin University)

Practical Gauss-Newton Optimisation for Deep Learning
Aleksandar Botev (University College London) · (None) · David Barber (University College London)

Estimating individual treatment effect: generalization bounds and algorithms
Uri Shalit (NYU) · Fredrik D Johansson (MIT) · David Sontag (Massachusetts Institute of Technology)

Online Multiview Learning: Dropping Convexity for Better Efficiency
Zhehui Chen (Georgia Institute of Technology) · Lin Yang (Johns Hopkins) · Chris Junchi Li (Princeton University) · Tuo Zhao (Georgia Institute of Technology)

Conditional Image Synthesis with Auxiliary Classifier GANs
Augustus Odena (Google Brain) · Christopher Olah (Google Brain) · Jon Shlens (Google Brain)

Variational Dropout Sparsifies Deep Neural Networks
Dmitry Molchanov (Skoltech) · Arsenii Ashukha (Moscow Institute of Physics and Technology) · Dmitry Vetrov (HSE)

Deep Bayesian Active Learning with Image Data
Yarin Gal (University of Cambridge) · Riashat Islam (McGill University) · Zoubin Ghahramani (University of Cambridge)

Active Learning for Cost-Sensitive Classification
Alekh Agarwal (Microsoft Research) · Akshay Krishnamurthy (UMass) · Tzu-Kuo Huang (Uber) · Hal Daumé III (University of Maryland) · John Langford (Microsoft Research)

Compressed Sensing using Generative Models
Ashish Bora (University of Texas at Austin) · Ajil Jalal (University of Texas at Austin) · Eric Price (UT-Austin) · Alexandros Dimakis (UT Austin)

Deriving Neural Architectures from Sequence and Graph Kernels
Tao Lei (MIT CSAIL) · Wengong Jin (MIT CSAIL) · Regina Barzilay (MIT CSAIL) · Tommi Jaakkola (MIT)

Variational Policy for Guiding Point Processes
Yichen Wang (Gatech) · Grady Williams (Georgia Tech) · Evangelos Theodorou (Georgia Tech) · Le Song (Georgia Institute of Technology)

Wasserstein Generative Adversarial Networks
Martin Arjovsky (New York University) · Soumith Chintala (Facebook) · Léon Bottou (Facebook)

Random Fourier Features for Kernel Ridge Regression: Approximation Bounds and Statistical Guarantees
Haim Avron (Tel Aviv University) · Michael Kapralov (EPFL) · Cameron Musco () · Christopher Musco () · Ameya Velingker () · Amir Zandieh (EPFL)

Selective Inference for Sparse High-Order Interaction Models
Shinya Suzumura (Nagoya Institute of Technology) · Kazuya Nakagawa (Nagoya Institute of Technology) · (None) · Koji Tsuda (University of Tokyo) · Ichiro Takeuchi (Nagoya Institute of Technology)

Globally Induced Forest: A Prepruning Compression Scheme
Jean-Michel Begon (University of Liege) · Arnaud Joly (University of Liege) · Pierre Geurts (University of Liege)

On The Projection Operator to A Three-view Cardinality Constrained Set
Haichuan Yang (University of Rochester) · Shupeng Gui (University of Rochester) · Chuyang Ke (University of Rochester) · Daniel Stefankovic (University of Rochester) · Ryohei Fujimaki () · Ji Liu ()

Diameter-Based Active Learning
Christopher Tosh (University of California, San Diego) · Sanjoy Dasgupta (UCSD)

Nonparanormal Information Estimation
Shashank Singh (Carnegie Mellon University) · Barnabás Póczos (CMU)

Convolutional Sequence to Sequence Learning
(None) · Michael Auli (Facebook) · David Grangier (Facebook) · Denis Yarats (Facebook AI Research) · Yann Dauphin (Facebook AI Research)

Adaptive Neural Networks for Fast Test-Time Prediction
Tolga Bolukbasi (Boston University) · Joseph Wang (Amazon) · Ofer Dekel (Microsoft) · (None)

On Calibration of Modern Neural Networks
Chuan Guo (Cornell University) · Geoff Pleiss (Cornell University) · Yu Sun (Cornell University) · Kilian Weinberger (Cornell University)

Programming with a Differentiable Forth Interpreter
Matko Bošnjak (University College London) · Tim Rocktäschel () · Jason Naradowsky () · Sebastian Riedel (UCL)

Follow the Moving Leader in Deep Learning
Shuai Zheng (Hong Kong University of Science and Technology) · James Kwok (Hong Kong University of Science and Technology)

A Unified Maximum Likelihood Approach for Estimating Symmetric Properties of Discrete Distributions
Jayadev Acharya (Cornell University) · Hirakendu Das (Yahoo!) · Alon Orlitsky (UCSD) · Ananda Suresh (Google)

Second-Order Kernel Online Convex Optimization with Adaptive Sketching
Daniele Calandriello (INRIA Lille) · Michal Valko (Inria Lille - Nord Europe) · Alessandro Lazaric ()

Tunable Efficient Unitary Neural Networks (EUNN) and their application to RNNs
Li Jing (Massachusetts Institute of Technology) · Yichen Shen (MIT) · Tena Dubcek (MIT) · John Peurifoy (MIT) · Scott Skirlo (MIT) · Yann LeCun (New York University) · Max Tegmark (MIT) · Marin Solja\v{c}i\'{c} (MIT)

Sequence Tutor: Conservative fine-tuning of sequence generation models with 057 003 KL-control
Natasha Jaques (Massachusetts Institute of Technology) · Shixiang Gu (Cambridge) · Dzmitry Bahdanau (Université de Montréal) · Jose Hernandez-Lobato

(University of Cambridge) · Rich Turner (University of Cambridge) · Douglas Eck (Google Brain)

On Relaxing Determinism in Arithmetic Circuits
Arthur Choi (UCLA) · Adnan Darwiche (UCLA)

Controllable Text Generation
(None) · Zichao Yang () · Xiaodan Liang (Carnegie Mellon University) · Ruslan Salakhutdinov (Carnegie Mellen University) · Eric Xing (Carnegie Mellon University)

Latent LSTM Allocation: Joint clustering and non-linear dynamic modeling of sequence data
Manzil Zaheer (CMU) · Amr Ahmed (Google) · Alex Smola (Amazon)

Recursive Partitioning for Personalization using Observational Data
Nathan Kallus (Cornell University)

Active Learning for Top-K Rank Aggregation from Noisy Comparisons
Soheil Mohajer (University of Minnesota) · Changho Suh (KAIST) · Adel Elmahdy (University of Minnesota)

Spectral Learning from a Single Trajectory under Finite-State Policies
Borja Balle (University of Lancaster) · Odalric Maillard ()

Learning to Align the Source Code to the Compiled Object Code
Dor Levy (Tel Aviv University) · Lior Wolf (Facebook AI Research and Tel Aviv University)

Zero-Shot Task Generalization with Multi-Task Deep Reinforcement Learning
Junhyuk Oh (University of Michigan) · Satinder Singh (University of Michigan) · Honglak Lee (Google / U. Michigan) · Pushmeet Kohli (Microsoft Research)

Multi-Class Optimal Margin Distribution Machine
Teng Zhang (Nanjing University) · Zhi-Hua Zhou (Nanjing University)

Bottleneck Conditional Density Estimation
Rui Shu (Stanford University) · Hung Bui (Adobe Research) · Mohammad Ghavamzadeh (Adobe)

A Divergence Bound for Hybrids of MCMC and Variational Inference and an Application to Langevin Dynamics and SGVI
Justin Domke (University of Massachusetts, Amherst)

Visualizing and Understanding Multilayer Perceptron Models: A Case Study in Speech Processing
Tasha Nagamine (Columbia University) · Nima Mesgarani (Columbia University)

Capacity rationed diffusions for speed and locality.
Satish Rao (UC Berkeley) · Di Wang () · Monika Henzinger () · Kimon Fountoulakis (University of California Berkeley and International Computer Science Institute) · Michael Mahoney (UC Berkeley)

Robust Structured Estimation with Single-Index Models
Sheng Chen (University of Minnesota) · Arindam Banerjee (University of Minnesota) · Sreangsu Acharyya (Microsoft Research India)

Stochastic Gradient MCMC Methods for Hidden Markov Models
Yi-An Ma (University of Washington) · Nick Foti (University of Washington) · Emily Fox (University of Washington)

Parseval Networks: Improving Robustness to Adversarial Examples
Moustapha Cisse (Facebook AI Research) · Piotr Bojanowski (Facebook) · Edouard Grave (Facebook AI Research) · Yann Dauphin (Facebook AI Research) · Nicolas Usunier (Facebook AI Research)

“Convex Until Proven Guilty”: Dimension-Free Acceleration of Gradient Descent on Non-Convex Functions
Oliver Hinder (Stanford) · Aaron Sidford (Stanford) · John Duchi (Stanford University) · Yair Carmon (Stanford)

Theoretical Properties for Neural Networks with Weight Matrices of Low Displacement Rank
Liang Zhao (The City University of New York) · Siyu Liao () · Yanzhi Wang () · Zhe Li (Syracuse University) · Jian Tang (Syracuse University) · Bo Yuan (City University of New York)

An Efficient, Sparsity-Preserving, Online Algorithm for Low-Rank Approximation
David Anderson (UC Berkeley) · Ming Gu ()

Improved Variational Autoencoders for Text Modeling using Dilated Convolutions
Zichao Yang () · Taylor Berg-Kirkpatrick () · (None) · Ruslan Salakhutdinov (Carnegie Mellen University)

Input Convex Neural Networks
Brandon Amos (Carnegie Mellon University) · Lei Xu (Carnegie Mellon University) · Zico Kolter (Carnegie Mellon University)

End-to-End Learning for Structured Prediction Energy Networks
David Belanger (UMass Amherst) · Bishan Yang (Carnegie Mellon University) · Andrew McCallum (UMass Amherst)

Convergence Analysis of Proximal Gradient with Momentum for Nonconvex Optimization
Qunwei Li (Syracuse University) · Yi Zhou (Syracuse University) · Yingbin Liang () · Pramod Varshney ()

Reinforcement Learning with Deep Energy-Based Policies
Tuomas Haarnoja (UC Berkeley) · Haoran Tang (UC Berkeley) · Pieter Abbeel (OpenAI / Berkeley) · Sergey Levine (Berkeley)

Count-Based Exploration with Neural Density Models
Georg Ostrovski (Google DeepMind) · Marc Bellemare (DeepMind) · Aäron van den Oord (Google) · Remi Munos (Google DeepMind)

Probabilistic Submodular Maximization in Sub-Linear Time
serban Stan (Yale) · morteza Zadimoghaddam (Google) · Andreas Krause (ETH Zurich) · Amin Karbasi (Yale)

On the Expressive Power of Deep Neural Networks
Maithra Raghu (Cornell University) · Ben Poole (Stanford University) · Surya Ganguli (Stanford) · Jon Kleinberg (Cornell University) · Jascha Sohl-Dickstein (Google Brain)

Neural Optimizer Search using Reinforcement Learning
Barret Zoph (Google) · Quoc Le (Google Brain) · Irwan Bello (Google) · Vijay Vasudevan (Google)

World of Bits: An Open-Domain Platform for Web-Based Agents
Tianlin Shi (Stanford University) · Andrej Karpathy (OpenAI) · Linxi Fan () · Jonathan Hernandez () · Percy Liang (Stanford University)

OptNet: Differentiable Optimization as a Layer in Neural Networks
Brandon Amos (Carnegie Mellon University) · Zico Kolter (Carnegie Mellon University)

Deep Decentralized Multi-task Multi-Agent Reinforcement Learning under Partial Observability
Shayegan Omidshafiei (MIT) · Jason Pazis (MIT) · Chris Amato (Northeastern University) · Jonathan How (MIT) · John L Vian (The Boeing Company)

Interactive Learning from Policy-Dependent Human Feedback
James MacGlashan (Cogitai) · Mark Ho (Brown University) · Robert Loftin (North Carolina State University) · Bei Peng (Washington State University) · Guan Wang (Brown University) · David Roberts (North Carolina State University) · Matthew Taylor (Washington State University) · Michael Littman (Brown University)

Differentially Private Chi-squared Test by Unit Circle Mechanism
Kazuya Kakizaki (NEC) · Jun Sakuma (University of Tsukuba) · Kazuto Fukuchi (University of Tsukuba)

Constrained Policy Optimization
Joshua Achiam (UC Berkeley) · David Held (UC Berkeley) · Aviv Tamar (UC Berkeley) · Pieter Abbeel (OpenAI / Berkeley)

Developing Bug-Free Machine Learning Systems With Formal Mathematics
Daniel Selsam (Stanford University) · David L Dill (Stanford University) · Percy Liang (Stanford University)

Axiomatic Attribution for Deep Networks
Ankur Taly (Google Inc.) · Qiqi Yan (Google Inc.) · Mukund Sundararajan (Google Inc.)

Gradient Coding: Avoiding Stragglers in Distributed Learning
Rashish Tandon (University of Texas at Austin) · Qi Lei (University of Texas at Austin) · Alexandros Dimakis (UT Austin) · NIKOS KARAMPATZIAKIS (Microsoft)

Learning Hierarchical Features from Generative Models
Shengjia Zhao (Stanford University) · Jiaming Song (Stanford University) · Stefano Ermon (Stanford University)

Combining Model-Based and Model-Free Updates for Trajectory-Centric Reinforcement Learning
Yevgen Chebotar (University of Southern California) · Karol Hausman (University of Southern California) · Marvin Zhang (UC Berkeley) · Gaurav Sukhatme (University of Southern California) · Stefan Schaal () · Sergey Levine (Berkeley)

Generalization and Equilibrium in Generative Adversarial Nets (GANs)
Sanjeev Arora (Princeton University) · Rong Ge (Duke University) · Yingyu Liang (Princeton University) · Tengyu Ma (Princeton University) · Yi Zhang (Princeton University)

Data-Efficient Policy Evaluation Through Behavior Policy Search
Josiah Hanna (University of Texas at Austin) · Philip Thomas (CMU) · Peter Stone (University of Texas at Austin) · Scott Niekum (University of Texas at Austin)

Stochastic Adaptive Quasi-Newton Methods for Minimizing Expected Values
Wenbo Gao (Columbia University) · Donald Goldfarb (Columbia University) · Chaoxu Zhou (Columbia University)

Fake News Mitigation via Point Process Based Intervention
Mehrdad Farajtabar (Georgia Tech) · Jiachen Yang (Georgia Institute of Technology) · Xiaojing Ye (Georgia State University) · Huan Xu (Georgia Tech) · Shuang Li () · Rakshit Trivedi (Georgia Institute of Technology) · Elias Khalil (Georgia Tech) · Le Song (Georgia Institute of Technology) · Hongyuan Zha (Georgia Institute of Technology)

Iterative Machine Teaching
Weiyang Liu (Georgia Tech) · Bo Dai (Georgia Tech) · Le Song (Georgia Institute of Technology)

Grammar Variational Autoencoder
Matt J. Kusner (Alan Turing Institute) · Brooks Paige (Alan Turing Institute) · Jose Hernandez-Lobato (University of Cambridge)

Collect at Once, Use Effectively: Making Non-interactive Locally Private Learning Possible
Kai Zheng (Peking University) · Wenlong Mou (Peking University) · Liwei Wang (Peking University)

Bayesian Sparsity for Intractable Undirected Models
John Ingraham (Harvard University) · Debora Marks (Harvard Medical School)

Reduced Space and Faster Convergence in Imperfect-Information Games via Pruning
Noam Brown (Carnegie Mellon University) · Tuomas Sandholm (Carnegie Mellon University)

Exploiting Strong Convexity from Data with Primal-Dual First-Order Algorithms
Jialei Wang (University of Chicago) · Lin Xiao (Microsoft Research)

Doubly Greedy Primal-Dual Coordinate Descent for Sparse Empirical Risk Minimization
Qi Lei (University of Texas at Austin) · Ian Yen (Carnegie Mellon University) · Chao-Yuan Wu (UT Austin) · Inderjit Dhillon (UT Austin & Amazon) · Pradeep Ravikumar (Carnegie Mellon University)

Emulating the Expert: Inverse Optimization through Online Learning
Sebastian Pokutta (Georgia Tech) · Andreas Bärmann (FAU Erlangen-Nürnberg) · Oskar Schneider ()

Online and Linear-Time Attention by Enforcing Monotonic Alignments
Colin Raffel (Google) · Thang Luong (Google Brain) · Peter Liu (Google) · Ron Weiss (Google Brain) · Douglas Eck (Google Brain)

The Latent Feature Lasso
Ian Yen (Carnegie Mellon University) · Wei-Chen Li (National Taiwan University) · Arun Suggala (Carnegie Mellon University) · Sung-En Chang (National Taiwan University) · Pradeep Ravikumar (Carnegie Mellon University) · Shou-De Lin (National Taiwan University)

Neural Audio Synthesis of Musical Notes with WaveNet Autoencoders
Cinjon Resnick (Google Brain) · Adam Roberts (Google Brain) · Jesse Engel (Google Brain) · Douglas Eck (Google Brain) · Sander Dieleman (DeepMind) · Karen Simonyan (DeepMind) · Mohammad Norouzi (Google)

Risk bounds for transferring representations with and without fine-tuning
Daniel McNamara (Australian National University and Data61) · Nina Balcan (Carnegie Mellon University)

Gradient Boosted Decision Trees for High Dimensional Sparse Output
Si Si (google research) · Huan Zhang (UC Davis) · Sathiya Keerthi (Microsoft) · Dhruv Mahajan (Facebook) · Inderjit Dhillon (UT Austin & Amazon) · Cho-Jui Hsieh (University of California)

Forest-type Regression with General Losses and Robust Forest
Hanbo Li (UC San Diego) · Andrew Martin (Zillow)

Counterfactual Data-Fusion for Online Reinforcement Learners
Andrew Forney (UCLA) · Elias Bareinboim (Purdue) · Judea Pearl (UCLA)

Efficient Optimization for Connected Subgraph Detection
Cem Aksoylar () · Orecchia Lorenzo () · (None)

A Closer Look at Memorization in Deep Networks
David Krueger (MILA) · Yoshua Bengio (U. Montreal) · Stanislaw Jastrzebsk () · Maxinder Kanwal () · Nicolas Ballas () · Asja Fischer () · Emmanuel Bengio () · Devansh Arpit () · Tegan Maharaj () · Aaron Courville (University of Montreal) · Simon Lacoste-Julien (University of Montreal)

Learning Gradient Descent: Better Generalization and Longer Horizons
Kaifeng Lv (Tsinghua University) · Shunhua Jiang (Tsinghua University) · Jian Li (IIIS)

Learning Deep Latent Gaussian Models with Markov Chain Monte Carlo
Matthew Hoffman (Google Research)

On Approximation Guarantees for Greedy Low Rank Optimization
Rajiv Khanna (UT Austin) · Ethan R. Elenberg (The University of Texas at Austin) · Alexandros Dimakis (UT Austin) · Sahand Negahban (YALE)

The Sample Complexity of Online One-Class Collaborative Filtering
Reinhard Heckel (UC Berkeley) · Kannan Ramchandran (UC Berkeley)

Algebraic Variety Models for High-Rank Matrix Completion
Greg Ongie (University of Michigan) · Laura Balzano (University of Michigan) · Rebecca Willett (UW Madison) · Robert Nowak (University of Wisconsion-Madison)

Learning Algorithms for Active Learning
Philip Bachman (Maluuba) · Alessandro Sordoni (Microsoft Maluuba) · Adam Trischler (Maluuba)

Maximum Selection and Ranking under Noisy Comparisons
Moein Falahatgar () · Alon Orlitsky (UCSD) · Venkatadheeraj Pichapati (University of California San Diego) · Ananda Suresh (Google Research)

Know-Evolve: Deep Learning for Temporal Reasoning in Dynamic Knowledge Graphs
Rakshit Trivedi (Georgia Institute of Technology) · Hajun Dai (Georgia Tech) · Yichen Wang (Gatech) · Le Song (Georgia Institute of Technology)

Deep IV: A Flexible Approach for Counterfactual Prediction
Greg Lewis (Microsoft Research) · Matt Taddy (MICROSOFT) · Jason Hartford (University of British Columbia) · Kevin Leyton-Brown ()

Variants of RMSProp and Adagrad with Logarithmic Regret Bounds
Mahesh Chandra Mukkamala (Universität des Saarlandes) · Matthias Hein (Saarland University)

Estimating the unseen from multiple populations
Aditi Raghunathan () · Greg Valiant () · James Zou (Stanford)

Stochastic DCA for the large-sum of non-convex functions problem. Application to group variables selection in multiclass logistic regression
Hoai Le Thi (University of Lorraine) · Duy Phan (Universite de Lorraine) · (None) · Bach Tran (University of Lorraine)

Language Modeling with Gated Convolutional Networks
Yann Dauphin (Facebook AI Research) · Angela Fan (Facebook AI Research) · Michael Auli (Facebook) · David Grangier (Facebook)

Device Placement Optimization with Reinforcement Learning
(None) · Hieu Pham (Google) · Quoc Le (Google Brain) · Mohammad Norouzi (Google) · Samy Bengio (Google Brain) · benoit steiner (Google) · Yuefeng Zhou (Google) · Naveen Kumar (Google) · Rasmus Larsen (Google) · Jeff Dean (Google)

Learning Sleep Stages from Radio Signals: A Deep Adversarial Architecture
Mingmin Zhao (MIT) · Shichao Yue (MIT) · Dina Katabi (MIT) · Tommi Jaakkola (MIT) · Matt Bianchi (Massachusetts General Hospital)

Stochastic Bouncy Particle Sampler
Ari Pakman (Columbia University) · Dar Gilboa (Columbia University) · David Carlson (Duke University) · Liam Paninski ()

Dissipativity Theory for Nesterov's Accelerated Method
Bin Hu (University of Wisconsin) · Laurent Lessard ()