

- 1, TITLE: AREs and MaRS Adversarial and MMD-Minimizing Regression for SDEs  
<http://proceedings.mlr.press/v97/abbati19a.html>  
AUTHORS: Gabriele Abbati, Philippe Wenk, Michael A. Osborne, Andreas Krause, Bernhard Schölkopf, Stefan Bauer  
HIGHLIGHT: In this paper, we propose a novel, probabilistic model for estimating the drift and diffusion given noisy observations of the underlying stochastic system.
- 2, TITLE: Dynamic Weights in Multi-Objective Deep Reinforcement Learning  
<http://proceedings.mlr.press/v97/abels19a.html>  
AUTHORS: Axel Abels, Diederik Roijers, Tom Lenaerts, Ann Nowé, Denis Steckelmacher  
HIGHLIGHT: We generalize across weight changes and high-dimensional inputs by proposing a multi-objective Q-network whose outputs are conditioned on the relative importance of objectives and we introduce Diverse Experience Replay (DER) to counter the inherent non-stationarity of the Dynamic Weights setting.
- 3, TITLE: MixHop: Higher-Order Graph Convolutional Architectures via Sparsified Neighborhood Mixing  
<http://proceedings.mlr.press/v97/abu-el-haija19a.html>  
AUTHORS: Sami Abu-El-Haija, Bryan Perozzi, Amol Kapoor, Nazanin Alipourfard, Kristina Lerman, Hrayr Harutyunyan, Greg Ver Steeg, Aram Galstyan  
HIGHLIGHT: To address this weakness, we propose a new model, MixHop, that can learn these relationships, including difference operators, by repeatedly mixing feature representations of neighbors at various distances.
- 4, TITLE: Communication-Constrained Inference and the Role of Shared Randomness  
<http://proceedings.mlr.press/v97/acharya19a.html>  
AUTHORS: Jayadev Acharya, Clement Canonne, Himanshu Tyagi  
HIGHLIGHT: We propose a general purpose simulate-and-infer strategy that uses only private-coin communication protocols and is sample-optimal for distribution learning.
- 5, TITLE: Distributed Learning with Sublinear Communication  
<http://proceedings.mlr.press/v97/acharya19b.html>  
AUTHORS: Jayadev Acharya, Chris De Sa, Dylan Foster, Karthik Sridharan  
HIGHLIGHT: Our main result is that by slightly relaxing the standard boundedness assumptions for linear models, we can obtain distributed algorithms that enjoy optimal error with communication logarithmic in dimension.
- 6, TITLE: Communication Complexity in Locally Private Distribution Estimation and Heavy Hitters  
<http://proceedings.mlr.press/v97/acharya19c.html>  
AUTHORS: Jayadev Acharya, Ziteng Sun  
HIGHLIGHT: We propose a sample-optimal  $\epsilon$ -locally differentially private (LDP) scheme for distribution estimation, where each user communicates one bit, and requires no public randomness.
- 7, TITLE: Learning Models from Data with Measurement Error: Tackling Underreporting  
<http://proceedings.mlr.press/v97/adams19a.html>  
AUTHORS: Roy Adams, Yuelong Ji, Xiaobin Wang, Suchi Saria  
HIGHLIGHT: In this paper we present a method for estimating the distribution of an outcome given a binary exposure that is subject to underreporting.  
As studies based on observational data are increasingly used to inform decisions with real-world impact, it is critical that we develop a robust set of techniques for analyzing and adjusting for these biases.
- 8, TITLE: TibGM: A Transferable and Information-Based Graphical Model Approach for Reinforcement Learning  
<http://proceedings.mlr.press/v97/adel19a.html>  
AUTHORS: Tameem Adel, Adrian Weller  
HIGHLIGHT: Here we propose a flexible GM-based RL framework which leverages efficient inference procedures to enhance generalisation and transfer power.

- 9, TITLE: PAC Learnability of Node Functions in Networked Dynamical Systems  
<http://proceedings.mlr.press/v97/adiga19a.html>  
AUTHORS: Abhijin Adiga, Chris J Kuhlman, Madhav Marathe, S Ravi, Anil Vullikanti  
HIGHLIGHT: We consider the PAC learnability of the local functions at the vertices of a discrete networked dynamical system, assuming that the underlying network is known.
- 10, TITLE: Static Automatic Batching In TensorFlow  
<http://proceedings.mlr.press/v97/agarwal19a.html>  
AUTHORS: Ashish Agarwal  
HIGHLIGHT: To address this we extend TensorFlow with pfor, a parallel-for loop optimized using static loop vectorization.
- 11, TITLE: Efficient Full-Matrix Adaptive Regularization  
<http://proceedings.mlr.press/v97/agarwal19b.html>  
AUTHORS: Naman Agarwal, Brian Bullins, Xinyi Chen, Elad Hazan, Karan Singh, Cyril Zhang, Yi Zhang  
HIGHLIGHT: We show how to modify full-matrix adaptive regularization in order to make it practical and effective.
- 12, TITLE: Online Control with Adversarial Disturbances  
<http://proceedings.mlr.press/v97/agarwal19c.html>  
AUTHORS: Naman Agarwal, Brian Bullins, Elad Hazan, Sham Kakade, Karan Singh  
HIGHLIGHT: We present an efficient algorithm that achieves nearly-tight regret bounds in this setting.
- 13, TITLE: Fair Regression: Quantitative Definitions and Reduction-Based Algorithms  
<http://proceedings.mlr.press/v97/agarwal19d.html>  
AUTHORS: Alekh Agarwal, Miroslav Dudik, Zhiwei Steven Wu  
HIGHLIGHT: In this paper, we study the prediction of a real-valued target, such as a risk score or recidivism rate, while guaranteeing a quantitative notion of fairness with respect to a protected attribute such as gender or race.
- 14, TITLE: Learning to Generalize from Sparse and Underspecified Rewards  
<http://proceedings.mlr.press/v97/agarwal19e.html>  
AUTHORS: Rishabh Agarwal, Chen Liang, Dale Schuurmans, Mohammad Norouzi  
HIGHLIGHT: We propose Meta Reward Learning (MeRL) to construct an auxiliary reward function that provides more refined feedback for learning.
- 15, TITLE: The Kernel Interaction Trick: Fast Bayesian Discovery of Pairwise Interactions in High Dimensions  
<http://proceedings.mlr.press/v97/agrawal19a.html>  
AUTHORS: Raj Agrawal, Brian Trippe, Jonathan Huggins, Tamara Broderick  
HIGHLIGHT: Our key insight is that many hierarchical models of practical interest admit a Gaussian process representation such that rather than maintaining a posterior over all  $O(p^2)$  interactions, we need only maintain a vector of  $O(p)$  kernel hyper-parameters.
- 16, TITLE: Understanding the Impact of Entropy on Policy Optimization  
<http://proceedings.mlr.press/v97/ahmed19a.html>  
AUTHORS: Zafarali Ahmed, Nicolas Le Roux, Mohammad Norouzi, Dale Schuurmans  
HIGHLIGHT: In this work, we analyze this claim using new visualizations of the optimization landscape based on randomly perturbing the loss function.
- 17, TITLE: Fairwashing: the risk of rationalization  
<http://proceedings.mlr.press/v97/aivodji19a.html>  
AUTHORS: Ulrich Aivodji, Hiromi Arai, Olivier Fortineau, S?bastien Gambis, Satoshi Hara, Alain Tapp  
HIGHLIGHT: Our solution, LaundryML, is based on a regularized rule list enumeration algorithm whose objective is to search for fair rule lists approximating an unfair black-box model.

18, TITLE: Adaptive Stochastic Natural Gradient Method for One-Shot Neural Architecture Search  
<http://proceedings.mlr.press/v97/akimoto19a.html>  
AUTHORS: Youhei Akimoto, Shinichi Shirakawa, Nozomu Yoshinari, Kento Uchida, Shota Saito, Kouhei Nishida  
HIGHLIGHT: We propose a stochastic natural gradient method with an adaptive step-size mechanism built upon our theoretical investigation (robust).

19, TITLE: Projections for Approximate Policy Iteration Algorithms  
<http://proceedings.mlr.press/v97/akrou19a.html>  
AUTHORS: Riad Akrou, Joni Pajarinen, Jan Peters, Gerhard Neumann  
HIGHLIGHT: In this paper, we propose to improve over such solutions by introducing a set of projections that transform the constrained problem into an unconstrained one which is then solved by standard gradient descent.

20, TITLE: Validating Causal Inference Models via Influence Functions  
<http://proceedings.mlr.press/v97/alaa19a.html>  
AUTHORS: Ahmed Alaa, Mihaela Van Der Schaar  
HIGHLIGHT: In this paper, we use influence functions  $\{—\}$  the functional derivatives of a loss function  $\{—\}$  to develop a model validation procedure that estimates the estimation error of causal inference methods.

21, TITLE: Multi-objective training of Generative Adversarial Networks with multiple discriminators  
<http://proceedings.mlr.press/v97/albuquerque19a.html>  
AUTHORS: Isabela Albuquerque, Joao Monteiro, Thang Doan, Breandan Considine, Tiago Falk, Ioannis Mitliagkas  
HIGHLIGHT: In this work, we revisit the multiple-discriminator setting by framing the simultaneous minimization of losses provided by different models as a multi-objective optimization problem.

22, TITLE: Graph Element Networks: adaptive, structured computation and memory  
<http://proceedings.mlr.press/v97/alet19a.html>  
AUTHORS: Ferran Alet, Adarsh Keshav Jeewajee, Maria Bauza Villalonga, Alberto Rodriguez, Tomas Lozano-Perez, Leslie Kaelbling  
HIGHLIGHT: We explore the use of graph neural networks (GNNs) to model spatial processes in which there is no a priori graphical structure.

23, TITLE: Analogies Explained: Towards Understanding Word Embeddings  
<http://proceedings.mlr.press/v97/allen19a.html>  
AUTHORS: Carl Allen, Timothy Hospedales  
HIGHLIGHT: We derive a probabilistically grounded definition of paraphrasing that we re-interpret as word transformation, a mathematical description of “ $\$w_x\$$  is to  $\$w_y\$$ ”.

24, TITLE: Infinite Mixture Prototypes for Few-shot Learning  
<http://proceedings.mlr.press/v97/allen19b.html>  
AUTHORS: Kelsey Allen, Evan Shelhamer, Hanul Shin, Joshua Tenenbaum  
HIGHLIGHT: We propose infinite mixture prototypes to adaptively represent both simple and complex data distributions for few-shot learning.

25, TITLE: A Convergence Theory for Deep Learning via Over-Parameterization  
<http://proceedings.mlr.press/v97/allen-zhu19a.html>  
AUTHORS: Zeyuan Allen-Zhu, Yuanzhi Li, Zhao Song  
HIGHLIGHT: In this work, we prove simple algorithms such as stochastic gradient descent (SGD) can find Global Minima on the training objective of DNNs in Polynomial Time.

26, TITLE: Asynchronous Batch Bayesian Optimisation with Improved Local Penalisation

<http://proceedings.mlr.press/v97/alvi19a.html>

AUTHORS: Ahsan Alvi, Binxin Ru, Jan-Peter Calliess, Stephen Roberts, Michael A. Osborne  
HIGHLIGHT: We address this problem by developing an approach, Penalising Locally for Asynchronous Bayesian Optimisation on K Workers (PLAyBOOK), for asynchronous parallel BO.

27, TITLE: Bounding User Contributions: A Bias-Variance Trade-off in Differential Privacy

<http://proceedings.mlr.press/v97/amin19a.html>

AUTHORS: Kareem Amin, Alex Kulesza, Andres Munoz, Sergei Vassiltiskii  
HIGHLIGHT: Here, we characterize this trade-off for an empirical risk minimization setting, showing that in general there is a “sweet spot” that depends on measurable properties of the dataset, but that there is also a concrete cost to privacy that cannot be avoided simply by collecting more data.

28, TITLE: Explaining Deep Neural Networks with a Polynomial Time Algorithm for Shapley Value Approximation

<http://proceedings.mlr.press/v97/ancona19a.html>

AUTHORS: Marco Ancona, Cengiz Oztireli, Markus Gross  
HIGHLIGHT: In this work, by leveraging recent results on uncertainty propagation, we propose a novel, polynomial-time approximation of Shapley values in deep neural networks.

29, TITLE: Scaling Up Ordinal Embedding: A Landmark Approach

<http://proceedings.mlr.press/v97/anderton19a.html>

AUTHORS: Jesse Anderton, Javed Aslam  
HIGHLIGHT: We propose a novel landmark-based method as a partial solution.

30, TITLE: Sorting Out Lipschitz Function Approximation

<http://proceedings.mlr.press/v97/anil19a.html>

AUTHORS: Cem Anil, James Lucas, Roger Grosse  
HIGHLIGHT: Based on this, we propose to combine a gradient norm preserving activation function, GroupSort, with norm-constrained weight matrices.

31, TITLE: Sparse Multi-Channel Variational Autoencoder for the Joint Analysis of Heterogeneous Data

<http://proceedings.mlr.press/v97/antelmi19a.html>

AUTHORS: Luigi Antelmi, Nicholas Ayache, Philippe Robert, Marco Lorenzi  
HIGHLIGHT: To tackle this problem, in this work we extend the variational framework of VAE to bring parsimony and interpretability when jointly account for latent relationships across multiple channels.

32, TITLE: Unsupervised Label Noise Modeling and Loss Correction

<http://proceedings.mlr.press/v97/arazo19a.html>

AUTHORS: Eric Arazo, Diego Ortego, Paul Albert, Noel O’Connor, Kevin Mcguinness  
HIGHLIGHT: Specifically, we propose a beta mixture to estimate this probability and correct the loss by relying on the network prediction (the so-called bootstrapping loss).

33, TITLE: Fine-Grained Analysis of Optimization and Generalization for Overparameterized Two-Layer Neural Networks

<http://proceedings.mlr.press/v97/arora19a.html>

AUTHORS: Sanjeev Arora, Simon Du, Wei Hu, Zhiyuan Li, Ruosong Wang  
HIGHLIGHT: This paper analyzes training and generalization for a simple 2-layer ReLU net with random initialization, and provides the following improvements over recent works: (i) Using a tighter characterization of training speed than recent papers, an explanation for why training a neural net with random labels leads to slower training, as originally observed in [Zhang et al. ICLR’17]. (ii) Generalization bound independent of network size, using a data-dependent complexity measure.

34, TITLE: Distributed Weighted Matching via Randomized Composable Coresets

<http://proceedings.mlr.press/v97/assadi19a.html>

- AUTHORS: Sepehr Assadi, Mohammadhossein Bateni, Vahab Mirrokni  
HIGHLIGHT: In this paper, we develop a simple distributed algorithm for the problem on general graphs with approximation guarantee of  $2 + \epsilon$  that (nearly) matches that of the sequential greedy algorithm.
- 35, TITLE: Stochastic Gradient Push for Distributed Deep Learning  
<http://proceedings.mlr.press/v97/assran19a.html>  
AUTHORS: Mahmoud Assran, Nicolas Loizou, Nicolas Ballas, Mike Rabbat  
HIGHLIGHT: This paper studies Stochastic Gradient Push (SGP), which combines PushSum with stochastic gradient updates.
- 36, TITLE: Bayesian Optimization of Composite Functions  
<http://proceedings.mlr.press/v97/astudillo19a.html>  
AUTHORS: Raul Astudillo, Peter Frazier  
HIGHLIGHT: We consider optimization of composite objective functions, i.e., of the form  $f(x) = g(h(x))$ , where  $h$  is a black-box derivative-free expensive-to-evaluate function with vector-valued outputs, and  $g$  is a cheap-to-evaluate real-valued function.
- 37, TITLE: Linear-Complexity Data-Parallel Earth Mover's Distance Approximations  
<http://proceedings.mlr.press/v97/atasu19a.html>  
AUTHORS: Kubilay Atasu, Thomas Mittelholzer  
HIGHLIGHT: We propose novel approximation algorithms that overcome both of these limitations, yet still achieve linear time complexity.
- 38, TITLE: Benefits and Pitfalls of the Exponential Mechanism with Applications to Hilbert Spaces and Functional PCA  
<http://proceedings.mlr.press/v97/awan19a.html>  
AUTHORS: Jordan Awan, Ana Kenney, Matthew Reimherr, Aleksandra Slavkovic  
HIGHLIGHT: We study its extension to settings with summaries based on infinite dimensional outputs such as with functional data analysis, shape analysis, and nonparametric statistics.
- 39, TITLE: Feature Grouping as a Stochastic Regularizer for High-Dimensional Structured Data  
<http://proceedings.mlr.press/v97/aydore19a.html>  
AUTHORS: Sergul Aydore, Bertrand Thirion, Gael Varoquaux  
HIGHLIGHT: We propose a new regularizer specifically designed to leverage structure in the data in a way that can be applied efficiently to complex models.
- 40, TITLE: Beyond the Chinese Restaurant and Pitman-Yor processes: Statistical Models with double power-law behavior  
<http://proceedings.mlr.press/v97/ayed19a.html>  
AUTHORS: Fadhel Ayed, Juho Lee, Francois Caron  
HIGHLIGHT: In this paper, we introduce a class of completely random measures which are doubly regularly-varying.
- 41, TITLE: Scalable Fair Clustering  
<http://proceedings.mlr.press/v97/backurs19a.html>  
AUTHORS: Arturs Backurs, Piotr Indyk, Krzysztof Onak, Baruch Schieber, Ali Vakilian, Tal Wagner  
HIGHLIGHT: In this paper, we present a practical approximate fairlet decomposition algorithm that runs in nearly linear time.
- 42, TITLE: Entropic GANs meet VAEs: A Statistical Approach to Compute Sample Likelihoods in GANs  
<http://proceedings.mlr.press/v97/balaji19a.html>  
AUTHORS: Yogesh Balaji, Hamed Hassani, Rama Chellappa, Soheil Feizi  
HIGHLIGHT: In this work, we resolve this issue by constructing an explicit probability model that can be used to compute sample likelihood statistics in GANs.

- 43, TITLE: Provable Guarantees for Gradient-Based Meta-Learning  
<http://proceedings.mlr.press/v97/balcan19a.html>  
AUTHORS: Maria-Florina Balcan, Mikhail Khodak, Ameet Talwalkar  
HIGHLIGHT: We study the problem of meta-learning through the lens of online convex optimization, developing a meta-algorithm bridging the gap between popular gradient-based meta-learning and classical regularization-based multi-task transfer methods.
- 44, TITLE: Open-ended learning in symmetric zero-sum games  
<http://proceedings.mlr.press/v97/balduzzi19a.html>  
AUTHORS: David Balduzzi, Marta Garnelo, Yoram Bachrach, Wojciech Czarnecki, Julien Perolat, Max Jaderberg, Thore Graepel  
HIGHLIGHT: In this paper, we introduce a geometric framework for formulating agent objectives in zero-sum games, in order to construct adaptive sequences of objectives that yield open-ended learning.
- 45, TITLE: Concrete Autoencoders: Differentiable Feature Selection and Reconstruction  
<http://proceedings.mlr.press/v97/balin19a.html>  
AUTHORS: Muhammed Fatih Balin, Abubakar Abid, James Zou  
HIGHLIGHT: We introduce the concrete autoencoder, an end-to-end differentiable method for global feature selection, which efficiently identifies a subset of the most informative features and simultaneously learns a neural network to reconstruct the input data from the selected features.
- 46, TITLE: HOList: An Environment for Machine Learning of Higher Order Logic Theorem Proving  
<http://proceedings.mlr.press/v97/bansal19a.html>  
AUTHORS: Kshitij Bansal, Sarah Loos, Markus Rabe, Christian Szegedy, Stewart Wilcox  
HIGHLIGHT: We present an environment, benchmark, and deep learning driven automated theorem prover for higher-order logic.
- 47, TITLE: Structured agents for physical construction  
<http://proceedings.mlr.press/v97/bapst19a.html>  
AUTHORS: Victor Bapst, Alvaro Sanchez-Gonzalez, Carl Doersch, Kimberly Stachenfeld, Pushmeet Kohli, Peter Battaglia, Jessica Hamrick  
HIGHLIGHT: We examine how a range of deep reinforcement learning agents fare on these challenges, and introduce several new approaches which provide superior performance.
- 48, TITLE: Learning to Route in Similarity Graphs  
<http://proceedings.mlr.press/v97/baranchuk19a.html>  
AUTHORS: Dmitry Baranchuk, Dmitry Persiyonov, Anton Sinitin, Artem Babenko  
HIGHLIGHT: In this paper we propose to learn the routing function that overcomes local minima via incorporating information about the graph global structure.
- 49, TITLE: A Personalized Affective Memory Model for Improving Emotion Recognition  
<http://proceedings.mlr.press/v97/barros19a.html>  
AUTHORS: Pablo Barros, German Parisi, Stefan Wernter  
HIGHLIGHT: In this paper, we present a neural model based on a conditional adversarial autoencoder to learn how to represent and edit general emotion expressions.
- 50, TITLE: Scale-free adaptive planning for deterministic dynamics & discounted rewards  
<http://proceedings.mlr.press/v97/bartlett19a.html>  
AUTHORS: Peter Bartlett, Victor Gabillon, Jennifer Healey, Michal Valko  
HIGHLIGHT: We introduce PlaTypOOS, an adaptive, robust, and efficient alternative to the OLOP (open-loop optimistic planning) algorithm.

51, TITLE: Pareto Optimal Streaming Unsupervised Classification  
<http://proceedings.mlr.press/v97/basu19a.html>  
AUTHORS: Soumya Basu, Steven Gutstein, Brent Lance, Sanjay Shakkottai  
HIGHLIGHT: In this paper, we characterize the Pareto-optimal region of accuracy and arrival rate, and develop an algorithm that can operate at any point within this region.

52, TITLE: Categorical Feature Compression via Submodular Optimization  
<http://proceedings.mlr.press/v97/bateni19a.html>  
AUTHORS: Mohammadhossein Bateni, Lin Chen, Hossein Esfandiari, Thomas Fu, Vahab Mirrokni, Afshin Rostamizadeh  
HIGHLIGHT: To address this, we introduce a novel re-parametrization of the mutual information objective, which we prove is submodular, and also design a data structure to query the submodular function in amortized  $O(\log n)$  time (where  $n$  is the input vocabulary size).

53, TITLE: Noise2Self: Blind Denoising by Self-Supervision  
<http://proceedings.mlr.press/v97/batson19a.html>  
AUTHORS: Joshua Batson, Loic Royer  
HIGHLIGHT: We propose a general framework for denoising high-dimensional measurements which requires no prior on the signal, no estimate of the noise, and no clean training data.

54, TITLE: Efficient optimization of loops and limits with randomized telescoping sums  
<http://proceedings.mlr.press/v97/beatson19a.html>  
AUTHORS: Alex Beatson, Ryan P Adams  
HIGHLIGHT: We propose randomized telescope (RT) gradient estimators, which represent the objective as the sum of a telescoping series and sample linear combinations of terms to provide cheap unbiased gradient estimates.

55, TITLE: Recurrent Kalman Networks: Factorized Inference in High-Dimensional Deep Feature Spaces  
<http://proceedings.mlr.press/v97/becke19a.html>  
AUTHORS: Philipp Becker, Harit Pandya, Gregor Gebhardt, Cheng Zhao, C. James Taylor, Gerhard Neumann  
HIGHLIGHT: We propose a new deep approach to Kalman filtering which can be learned directly in an end-to-end manner using backpropagation without additional approximations.

56, TITLE: Switching Linear Dynamics for Variational Bayes Filtering  
<http://proceedings.mlr.press/v97/becke-ehmck19a.html>  
AUTHORS: Philip Becker-Ehmck, Jan Peters, Patrick Van Der Smagt  
HIGHLIGHT: Leveraging Bayesian inference, Variational Autoencoders and Concrete relaxations, we show how to learn a richer and more meaningful state space, e.g. encoding joint constraints and collisions with walls in a maze, from partial and high-dimensional observations.

57, TITLE: Active Learning for Probabilistic Structured Prediction of Cuts and Matchings  
<http://proceedings.mlr.press/v97/behpour19a.html>  
AUTHORS: Sima Behpour, Anqi Liu, Brian Ziebart  
HIGHLIGHT: We propose an adversarial approach for active learning with structured prediction domains that is tractable for cuts and matching.

58, TITLE: Invertible Residual Networks  
<http://proceedings.mlr.press/v97/behmann19a.html>  
AUTHORS: Jens Behrmann, Will Grathwohl, Ricky T. Q. Chen, David Duvenaud, Joern-Henrik Jacobsen  
HIGHLIGHT: To compute likelihoods, we introduce a tractable approximation to the Jacobian log-determinant of a residual block.

59, TITLE: Greedy Layerwise Learning Can Scale To ImageNet  
<http://proceedings.mlr.press/v97/belilovsky19a.html>  
AUTHORS: Eugene Belilovsky, Michael Eickenberg, Edouard Oyallon

**HIGHLIGHT:** Here we use 1-hidden layer learning problems to sequentially build deep networks layer by layer, which can inherit properties from shallow networks.

60, **TITLE:** Overcoming Multi-model Forgetting

<http://proceedings.mlr.press/v97/benyahia19a.html>

**AUTHORS:** Yassine Benyahia, Kaicheng Yu, Kamil Bennani Smires, Martin Jaggi, Anthony C. Davison, Mathieu Salzmann, Claudiu Musat

**HIGHLIGHT:** To overcome this, we introduce a statistically-justified weight plasticity loss that regularizes the learning of a model's shared parameters according to their importance for the previous models, and demonstrate its effectiveness when training two models sequentially and for neural architecture search.

61, **TITLE:** Optimal Kronecker-Sum Approximation of Real Time Recurrent Learning

<http://proceedings.mlr.press/v97/benzing19a.html>

**AUTHORS:** Frederik Benzing, Marcelo Matheus Gauy, Asier Mujika, Anders Martinsson, Angelika Steger

**HIGHLIGHT:** We present a new approximation algorithm of RTRL, Optimal Kronecker-Sum Approximation (OK).

62, **TITLE:** Adversarially Learned Representations for Information Obfuscation and Inference

<http://proceedings.mlr.press/v97/bertran19a.html>

**AUTHORS:** Martin Bertran, Natalia Martinez, Afroditi Papadaki, Qiang Qiu, Miguel Rodrigues, Galen Reeves, Guillermo Sapiro

**HIGHLIGHT:** In this work, we take an information theoretic approach that is implemented as an unconstrained adversarial game between Deep Neural Networks in a principled, data-driven manner.

63, **TITLE:** Bandit Multiclass Linear Classification: Efficient Algorithms for the Separable Case

<http://proceedings.mlr.press/v97/beygelzimer19a.html>

**AUTHORS:** Alina Beygelzimer, David Pal, Balazs Szorenyi, Devanathan Thiruvengatathari, Chen-Yu Wei, Chicheng Zhang

**HIGHLIGHT:** In this work, we take a first step towards this problem.

64, **TITLE:** Analyzing Federated Learning through an Adversarial Lens

<http://proceedings.mlr.press/v97/bhagoji19a.html>

**AUTHORS:** Arjun Nitin Bhagoji, Supriyo Chakraborty, Prateek Mittal, Seraphin Calo

**HIGHLIGHT:** In this work, we explore how the federated learning setting gives rise to a new threat, namely model poisoning, which differs from traditional data poisoning.

65, **TITLE:** Optimal Continuous DR-Submodular Maximization and Applications to Provable Mean Field Inference

<http://proceedings.mlr.press/v97/bian19a.html>

**AUTHORS:** Yatao Bian, Joachim Buhmann, Andreas Krause

**HIGHLIGHT:** In this work we propose provable mean field methods for probabilistic log-submodular models and its posterior agreement (PA) with strong approximation guarantees.

66, **TITLE:** More Efficient Off-Policy Evaluation through Regularized Targeted Learning

<http://proceedings.mlr.press/v97/bibaut19a.html>

**AUTHORS:** Aurelien Bibaut, Ivana Malenica, Nikos Vlassis, Mark Van Der Laan

**HIGHLIGHT:** In particular, we introduce a novel doubly-robust estimator for the OPE problem in RL, based on the Targeted Maximum Likelihood Estimation principle from the statistical causal inference literature.

67, **TITLE:** A Kernel Perspective for Regularizing Deep Neural Networks

<http://proceedings.mlr.press/v97/biotti19a.html>

**AUTHORS:** Alberto Biotti, Grégoire Mialon, Dexiong Chen, Julien Mairal

**HIGHLIGHT:** We propose a new point of view for regularizing deep neural networks by using the norm of a reproducing kernel Hilbert space (RKHS).



68, TITLE: Rethinking Lossy Compression: The Rate-Distortion-Perception Tradeoff  
<http://proceedings.mlr.press/v97/blau19a.html>  
AUTHORS: Yochai Blau, Tomer Michaeli  
HIGHLIGHT: In this paper, we adopt the mathematical definition of perceptual quality recently proposed by Blau & Michaeli (2018), and use it to study the three-way tradeoff between rate, distortion, and perception.

69, TITLE: Correlated bandits or: How to minimize mean-squared error online  
<http://proceedings.mlr.press/v97/boda19a.html>  
AUTHORS: Vinay Praneeth Boda, Prashanth L.A.  
HIGHLIGHT: Under a best-arm identification framework, we propose a successive rejects type algorithm and provide bounds on the probability of error in identifying the best arm.

70, TITLE: Adversarial Attacks on Node Embeddings via Graph Poisoning  
<http://proceedings.mlr.press/v97/bojchevski19a.html>  
AUTHORS: Aleksandar Bojchevski, Stephan G.lemnemann  
HIGHLIGHT: We provide the first adversarial vulnerability analysis on the widely used family of methods based on random walks.

71, TITLE: Online Variance Reduction with Mixtures  
<http://proceedings.mlr.press/v97/borsos19a.html>  
AUTHORS: Zoltan Borsos, Sebastian Curi, Kfir Yehuda Levy, Andreas Krause  
HIGHLIGHT: In this work, we propose a new framework for variance reduction that enables the use of mixtures over predefined sampling distributions, which can naturally encode prior knowledge about the data.

72, TITLE: Compositional Fairness Constraints for Graph Embeddings  
<http://proceedings.mlr.press/v97/bose19a.html>  
AUTHORS: Avishek Bose, William Hamilton  
HIGHLIGHT: Here, we introduce an adversarial framework to enforce fairness constraints on graph embeddings.

73, TITLE: Unreproducible Research is Reproducible  
<http://proceedings.mlr.press/v97/bouthillier19a.html>  
AUTHORS: Xavier Bouthillier, Cesar Laurent, Pascal Vincent  
HIGHLIGHT: This work is an attempt to promote the use of more rigorous and diversified methodologies.

74, TITLE: Blended Conditional Gradients  
<http://proceedings.mlr.press/v97/braun19a.html>  
AUTHORS: Gabor Braun, Sebastian Pokutta, Dan Tu, Stephen Wright  
HIGHLIGHT: We present a blended conditional gradient approach for minimizing a smooth convex function over a polytope  $P$ , combining the Frank-Wolfe algorithm (also called conditional gradient) with gradient-based steps, different from away steps and pairwise steps, but still achieving linear convergence for strongly convex functions, along with good practical performance.

75, TITLE: Coresets for Ordered Weighted Clustering  
<http://proceedings.mlr.press/v97/braverman19a.html>  
AUTHORS: Vladimir Braverman, Shaofeng H.-C. Jiang, Robert Krauthgamer, Xuan Wu  
HIGHLIGHT: Our main result is a construction of a simultaneous coreset of size  $O(d(k \log 2|X|))$  for Ordered  $k$ -Median.

76, TITLE: Target Tracking for Contextual Bandits: Application to Demand Side Management  
<http://proceedings.mlr.press/v97/bregere19a.html>  
AUTHORS: Margaux Br?g?re, Pierre Gaillard, Yannig Goude, Gilles Stoltz

HIGHLIGHT: We propose a contextual-bandit approach for demand side management by offering price incentives.

77, TITLE: Active Manifolds: A non-linear analogue to Active Subspaces

<http://proceedings.mlr.press/v97/bridges19a.html>

AUTHORS: Robert Bridges, Anthony Gruber, Christopher Felder, Miki Verma, Chelsey Hoff

HIGHLIGHT: We present an approach to analyze  $\mathcal{C}^1(\mathbb{R}^m)$  functions that addresses limitations present in the Active Subspaces (AS) method of Constantine et al. (2014; 2015).

78, TITLE: Conditioning by adaptive sampling for robust design

<http://proceedings.mlr.press/v97/brookes19a.html>

AUTHORS: David Brookes, Hahnbeom Park, Jennifer Listgarten

HIGHLIGHT: We present a method for design problems wherein the goal is to maximize or specify the value of one or more properties of interest (e.g. maximizing the fluorescence of a protein).

79, TITLE: Extrapolating Beyond Suboptimal Demonstrations via Inverse Reinforcement Learning from Observations

<http://proceedings.mlr.press/v97/brown19a.html>

AUTHORS: Daniel Brown, Wonjoon Goo, Prabhat Nagarajan, Scott Niekum

HIGHLIGHT: In this paper, we introduce a novel reward-learning-from-observation algorithm, Trajectory-ranked Reward EXtrapolation (T-REX), that extrapolates beyond a set of (approximately) ranked demonstrations in order to infer high-quality reward functions from a set of potentially poor demonstrations.

80, TITLE: Deep Counterfactual Regret Minimization

<http://proceedings.mlr.press/v97/brown19b.html>

AUTHORS: Noam Brown, Adam Lerer, Sam Gross, Tuomas Sandholm

HIGHLIGHT: This paper introduces Deep Counterfactual Regret Minimization, a form of CFR that obviates the need for abstraction by instead using deep neural networks to approximate the behavior of CFR in the full game.

81, TITLE: Understanding the Origins of Bias in Word Embeddings

<http://proceedings.mlr.press/v97/brunet19a.html>

AUTHORS: Marc-Etienne Brunet, Colleen Alkalay-Houlihan, Ashton Anderson, Richard Zemel

HIGHLIGHT: In this work we develop a technique to address this question.

82, TITLE: Low Latency Privacy Preserving Inference

<http://proceedings.mlr.press/v97/brutzkus19a.html>

AUTHORS: Alon Brutzkus, Ran Gilad-Bachrach, Oren Elisha

HIGHLIGHT: In this study we provide two solutions that address these limitations.

83, TITLE: Why do Larger Models Generalize Better? A Theoretical Perspective via the XOR Problem

<http://proceedings.mlr.press/v97/brutzkus19b.html>

AUTHORS: Alon Brutzkus, Amir Globerson

HIGHLIGHT: In this work, we provide theoretical and empirical evidence that, in certain cases, overparameterized convolutional networks generalize better than small networks because of an interplay between weight clustering and feature exploration at initialization.

84, TITLE: Adversarial examples from computational constraints

<http://proceedings.mlr.press/v97/bubeck19a.html>

AUTHORS: Sebastien Bubeck, Yin Tat Lee, Eric Price, Ilya Razenshteyn

HIGHLIGHT: Why are classifiers in high dimension vulnerable to "adversarial" perturbations? We show that it is likely not due to information theoretic limitations, but rather it could be due to computational constraints.

85, TITLE: Self-similar Epochs: Value in arrangement

<http://proceedings.mlr.press/v97/buchnik19a.html>

AUTHORS: Eliav Buchnik, Edith Cohen, Avinatan Hasidim, Yossi Matias

HIGHLIGHT: We hypothesize that the training can be more effective with self-similar arrangements that potentially allow each epoch to provide benefits of multiple ones.

86, TITLE: Learning Generative Models across Incomparable Spaces

<http://proceedings.mlr.press/v97/bunne19a.html>

AUTHORS: Charlotte Bunne, David Alvarez-Melis, Andreas Krause, Stefanie Jegelka

HIGHLIGHT: In this work, we propose an approach to learn generative models across such incomparable spaces, and demonstrate how to steer the learned distribution towards target properties.

87, TITLE: Rates of Convergence for Sparse Variational Gaussian Process Regression

<http://proceedings.mlr.press/v97/burt19a.html>

AUTHORS: David Burt, Carl Edward Rasmussen, Mark Van Der Wilk

HIGHLIGHT: We show that with high probability the KL divergence can be made arbitrarily small by growing  $M$  more slowly than  $N$ .

88, TITLE: What is the Effect of Importance Weighting in Deep Learning?

<http://proceedings.mlr.press/v97/byrd19a.html>

AUTHORS: Jonathon Byrd, Zachary Lipton

HIGHLIGHT: We present the surprising finding that while importance weighting impacts models early in training, its effect diminishes over successive epochs.

89, TITLE: A Quantitative Analysis of the Effect of Batch Normalization on Gradient Descent

<http://proceedings.mlr.press/v97/cai19a.html>

AUTHORS: Yongqiang Cai, Qianxiao Li, Zuowei Shen

HIGHLIGHT: In this paper, we provide such an analysis on the simple problem of ordinary least squares (OLS), where the precise dynamical properties of gradient descent (GD) is completely known, thus allowing us to isolate and compare the additional effects of BN.

90, TITLE: Accelerated Linear Convergence of Stochastic Momentum Methods in Wasserstein Distances

<http://proceedings.mlr.press/v97/can19a.html>

AUTHORS: Bugra Can, Mert Gurbuzbalaban, Lingjiong Zhu

HIGHLIGHT: For strongly convex problems, we show that the distribution of the iterates of AG converges with the accelerated  $O(\sqrt{\kappa} \log(1/\epsilon))$  linear rate to a ball of radius  $\epsilon$  centered at a unique invariant distribution in the 1-Wasserstein metric where  $\kappa$  is the condition number as long as the noise variance is smaller than an explicit upper bound we can provide.

91, TITLE: Active Embedding Search via Noisy Paired Comparisons

<http://proceedings.mlr.press/v97/canal19a.html>

AUTHORS: Gregory Canal, Andy Massimino, Mark Davenport, Christopher Rozell

HIGHLIGHT: In such tasks, queries can be extremely costly and subject to varying levels of response noise; thus, we aim to actively choose pairs that are most informative given the results of previous comparisons.

92, TITLE: Dynamic Learning with Frequent New Product Launches: A Sequential Multinomial Logit Bandit Problem

<http://proceedings.mlr.press/v97/cao19a.html>

AUTHORS: Junyu Cao, Wei Sun

HIGHLIGHT: For the offline version with known customers' preferences, we propose a polynomial-time algorithm and characterize the properties of the optimal tiered product recommendation.

93, TITLE: Competing Against Nash Equilibria in Adversarially Changing Zero-Sum Games

<http://proceedings.mlr.press/v97/cardoso19a.html>

AUTHORS: Adrian Rivera Cardoso, Jacob Abernethy, He Wang, Huan Xu

**HIGHLIGHT:** But when the payoff matrix evolves over time our goal is to find a sequential algorithm that can compete with, in a certain sense, the NE of the long-term-averaged payoff matrix.

94, **TITLE:** Automated Model Selection with Bayesian Quadrature

<http://proceedings.mlr.press/v97/chai19a.html>

**AUTHORS:** Henry Chai, Jean-Francois Ton, Michael A. Osborne, Roman Garnett

**HIGHLIGHT:** We present a novel technique for tailoring Bayesian quadrature (BQ) to model selection.

95, **TITLE:** Learning Action Representations for Reinforcement Learning

<http://proceedings.mlr.press/v97/chandak19a.html>

**AUTHORS:** Yash Chandak, Georgios Theodorou, James Kostas, Scott Jordan, Philip Thomas

**HIGHLIGHT:** We provide an algorithm to both learn and use action representations and provide conditions for its convergence.

96, **TITLE:** Dynamic Measurement Scheduling for Event Forecasting using Deep RL

<http://proceedings.mlr.press/v97/chang19a.html>

**AUTHORS:** Chun-Hao Chang, Mingjie Mai, Anna Goldenberg

**HIGHLIGHT:** We answer this question by deep reinforcement learning (RL) that jointly minimizes the measurement cost and maximizes predictive gain, by scheduling strategically-timed measurements.

97, **TITLE:** On Symmetric Losses for Learning from Corrupted Labels

<http://proceedings.mlr.press/v97/charoenphakdee19a.html>

**AUTHORS:** Nontawat Charoenphakdee, Jongyeong Lee, Masashi Sugiyama

**HIGHLIGHT:** This paper aims to provide a better understanding of a symmetric loss.

98, **TITLE:** Online learning with kernel losses

<http://proceedings.mlr.press/v97/chatterji19a.html>

**AUTHORS:** Niladri Chatterji, Aldo Pacchiano, Peter Bartlett

**HIGHLIGHT:** We present a generalization of the adversarial linear bandits framework, where the underlying losses are kernel functions (with an associated reproducing kernel Hilbert space) rather than linear functions.

99, **TITLE:** Neural Network Attributions: A Causal Perspective

<http://proceedings.mlr.press/v97/chattopadhyay19a.html>

**AUTHORS:** Aditya Chattopadhyay, Piyushi Manupriya, Anirban Sarkar, Vineeth N Balasubramanian

**HIGHLIGHT:** We propose a new attribution method for neural networks developed using ?rst principles of causality (to the best of our knowledge, the ?rst such).

100, **TITLE:** PAC Identification of Many Good Arms in Stochastic Multi-Armed Bandits

<http://proceedings.mlr.press/v97/chaudhuri19a.html>

**AUTHORS:** Arghya Roy Chaudhuri, Shivaram Kalyanakrishnan

**HIGHLIGHT:** We present a lower bound on the worst-case sample complexity for general  $k$ , and a fully sequential PAC algorithm, LUCB- $k$ - $m$ , which is more sample-efficient on easy instances.

101, **TITLE:** Nearest Neighbor and Kernel Survival Analysis: Nonasymptotic Error Bounds and Strong Consistency Rates

<http://proceedings.mlr.press/v97/chen19a.html>

**AUTHORS:** George Chen

**HIGHLIGHT:** We establish the first nonasymptotic error bounds for Kaplan-Meier-based nearest neighbor and kernel survival probability estimators where feature vectors reside in metric spaces.

102, **TITLE:** Stein Point Markov Chain Monte Carlo

<http://proceedings.mlr.press/v97/chen19b.html>

AUTHORS: Wilson Ye Chen, Alessandro Barp, Francois-Xavier Briol, Jackson Gorham, Mark Girolami, Lester Mackey, Chris Oates  
HIGHLIGHT: Stein Point Markov Chain Monte Carlo

103, TITLE: Particle Flow Bayes? Rule  
<http://proceedings.mlr.press/v97/chen19c.html>  
AUTHORS: Xinshi Chen, Hanjun Dai, Le Song  
HIGHLIGHT: We present a particle flow realization of Bayes' rule, where an ODE-based neural operator is used to transport particles from a prior to its posterior after a new observation.

104, TITLE: Proportionally Fair Clustering  
<http://proceedings.mlr.press/v97/chen19d.html>  
AUTHORS: Xingyu Chen, Brandon Fain, Liang Lyu, Kamesh Munagala  
HIGHLIGHT: We present and analyze algorithms to efficiently compute, optimize, and audit proportional solutions.

105, TITLE: Information-Theoretic Considerations in Batch Reinforcement Learning  
<http://proceedings.mlr.press/v97/chen19e.html>  
AUTHORS: Jinglin Chen, Nan Jiang  
HIGHLIGHT: In this paper, we revisit these assumptions and provide theoretical results towards answering the above questions, and make steps towards a deeper understanding of value-function approximation.

106, TITLE: Generative Adversarial User Model for Reinforcement Learning Based Recommendation System  
<http://proceedings.mlr.press/v97/chen19f.html>  
AUTHORS: Xinshi Chen, Shuang Li, Hui Li, Shaohua Jiang, Yuan Qi, Le Song  
HIGHLIGHT: In this paper, we propose a novel model-based reinforcement learning framework for recommendation systems, where we develop a generative adversarial network to imitate user behavior dynamics and learn her reward function.

107, TITLE: Understanding and Utilizing Deep Neural Networks Trained with Noisy Labels  
<http://proceedings.mlr.press/v97/chen19g.html>  
AUTHORS: Pengfei Chen, Ben Ben Liao, Guangyong Chen, Shengyu Zhang  
HIGHLIGHT: In this paper, we find that the test accuracy can be quantitatively characterized in terms of the noise ratio in datasets.

108, TITLE: A Gradual, Semi-Discrete Approach to Generative Network Training via Explicit Wasserstein Minimization  
<http://proceedings.mlr.press/v97/chen19h.html>  
AUTHORS: Yucheng Chen, Matus Telgarsky, Chao Zhang, Bolton Bailey, Daniel Hsu, Jian Peng  
HIGHLIGHT: This paper provides a simple procedure to fit generative networks to target distributions, with the goal of a small Wasserstein distance (or other optimal transport costs).

109, TITLE: Transferability vs. Discriminability: Batch Spectral Penalization for Adversarial Domain Adaptation  
<http://proceedings.mlr.press/v97/chen19i.html>  
AUTHORS: Xinyang Chen, Sinan Wang, Mingsheng Long, Jianmin Wang  
HIGHLIGHT: In this paper, a series of experiments based on spectral analysis of the feature representations have been conducted, revealing an unexpected deterioration of the discriminability while learning transferable features adversarially.

110, TITLE: Fast Incremental von Neumann Graph Entropy Computation: Theory, Algorithm, and Applications  
<http://proceedings.mlr.press/v97/chen19j.html>  
AUTHORS: Pin-Yu Chen, Lingfei Wu, Sijia Liu, Indika Rajapakse

**HIGHLIGHT:** In this paper, we propose a new computational framework, Fast Incremental von Neumann Graph EntRopy (FINGER), which approaches VNGE with a performance guarantee.

111, **TITLE:** Catalyst: Boosting Convex Katyusha for Non-Convex Problems with a Large Condition Number  
<http://proceedings.mlr.press/v97/chen19k.html>

**AUTHORS:** Zaiyi Chen, Yi Xu, Haoyuan Hu, Tianbao Yang

**HIGHLIGHT:** In this paper, we present a simple but non-trivial boosting of a state-of-the-art SVRG-type method for convex problems (namely Katyusha) to enjoy an improved complexity for solving non-convex problems with a large condition number (that is close to a convex function).

112, **TITLE:** Multivariate-Information Adversarial Ensemble for Scalable Joint Distribution Matching  
<http://proceedings.mlr.press/v97/chen19l.html>

**AUTHORS:** Ziliang Chen, Zhanfu Yang, Xiaoxi Wang, Xiaodan Liang, Xiaopeng Yan, Guanbin Li, Liang Lin

**HIGHLIGHT:** In this paper, we propose a domain-scalable DGM, i.e., MMI-ALI for  $\mathcal{M}$ -domain joint distribution matching.

113, **TITLE:** Robust Decision Trees Against Adversarial Examples

<http://proceedings.mlr.press/v97/chen19m.html>

**AUTHORS:** Hongge Chen, Huan Zhang, Duane Boning, Cho-Jui Hsieh

**HIGHLIGHT:** In this paper, we show that tree-based models are also vulnerable to adversarial examples and develop a novel algorithm to learn robust trees.

114, **TITLE:** RaFM: Rank-Aware Factorization Machines

<http://proceedings.mlr.press/v97/chen19n.html>

**AUTHORS:** Xiaoshuang Chen, Yin Zheng, Jiaying Wang, Wenye Ma, Junzhou Huang

**HIGHLIGHT:** Different from existing FM-based approaches which use a fixed rank for all features, this paper proposes a Rank-Aware FM (RaFM) model which adopts pairwise interactions from embeddings with different ranks.

115, **TITLE:** Control Regularization for Reduced Variance Reinforcement Learning

<http://proceedings.mlr.press/v97/cheng19a.html>

**AUTHORS:** Richard Cheng, Abhinav Verma, Gabor Orosz, Swarat Chaudhuri, Yisong Yue, Joel Burdick

**HIGHLIGHT:** Focusing on problems arising in continuous control, we propose a functional regularization approach to augmenting model-free RL.

116, **TITLE:** Predictor-Corrector Policy Optimization

<http://proceedings.mlr.press/v97/cheng19b.html>

**AUTHORS:** Ching-An Cheng, Xinyan Yan, Nathan Ratliff, Byron Boots

**HIGHLIGHT:** We present a predictor-corrector framework, called PicCoLO, that can transform a first-order model-free reinforcement or imitation learning algorithm into a new hybrid method that leverages predictive models to accelerate policy learning.

117, **TITLE:** Variational Inference for sparse network reconstruction from count data

<http://proceedings.mlr.press/v97/chiquet19a.html>

**AUTHORS:** Julien Chiquet, Stephane Robin, Mahendra Mariadassou

**HIGHLIGHT:** In this work, we consider instead a full-fledged probabilistic model with a latent layer where the counts follow Poisson distributions, conditional to latent (hidden) Gaussian correlated variables.

118, **TITLE:** Random Walks on Hypergraphs with Edge-Dependent Vertex Weights

<http://proceedings.mlr.press/v97/chitra19a.html>

**AUTHORS:** Uthsav Chitra, Benjamin Raphael

**HIGHLIGHT:** In this paper, we use random walks to develop a spectral theory for hypergraphs with edge-dependent vertex weights: hypergraphs where every vertex  $v$  has a weight  $\gamma_e(v)$  for each incident hyperedge  $e$  that describes the contribution of  $v$  to the hyperedge  $e$ .

- 119, TITLE: Neural Joint Source-Channel Coding  
<http://proceedings.mlr.press/v97/choi19a.html>  
AUTHORS: Kristy Choi, Kedar Tatwawadi, Aditya Grover, Tsachy Weissman, Stefano Ermon  
HIGHLIGHT: In this work, we propose to jointly learn the encoding and decoding processes using a new discrete variational autoencoder model.
- 120, TITLE: Beyond Backprop: Online Alternating Minimization with Auxiliary Variables  
<http://proceedings.mlr.press/v97/choromanska19a.html>  
AUTHORS: Anna Choromanska, Benjamin Cowen, Sadhana Kumaravel, Ronny Luss, Mattia Rigotti, Irina Rish, Paolo Diachille, Viatcheslav Gurev, Brian Kingsbury, Ravi Tejwani, Djallel Bouneffouf  
HIGHLIGHT: The main contribution of our work is a novel online (stochastic/mini-batch) alternating minimization (AM) approach for training deep neural networks, together with the first theoretical convergence guarantees for AM in stochastic settings and promising empirical results on a variety of architectures and datasets.
- 121, TITLE: Unifying Orthogonal Monte Carlo Methods  
<http://proceedings.mlr.press/v97/choromanski19a.html>  
AUTHORS: Krzysztof Choromanski, Mark Rowland, Wenyu Chen, Adrian Weller  
HIGHLIGHT: In this paper, we present a unifying perspective of many approximate methods by considering Givens transformations, propose new approximate methods based on this framework, and demonstrate the first statistical guarantees for families of approximate methods in kernel approximation.
- 122, TITLE: Probability Functional Descent: A Unifying Perspective on GANs, Variational Inference, and Reinforcement Learning  
<http://proceedings.mlr.press/v97/chu19a.html>  
AUTHORS: Casey Chu, Jose Blanchet, Peter Glynn  
HIGHLIGHT: The goal of this paper is to provide a unifying view of a wide range of problems of interest in machine learning by framing them as the minimization of functionals defined on the space of probability measures.
- 123, TITLE: MeanSum: A Neural Model for Unsupervised Multi-Document Abstractive Summarization  
<http://proceedings.mlr.press/v97/chu19b.html>  
AUTHORS: Eric Chu, Peter Liu  
HIGHLIGHT: In our work, we consider the setting where there are only documents (product or business reviews) with no summaries provided, and propose an end-to-end, neural model architecture to perform unsupervised abstractive summarization.  
Finally, we collect a ground-truth evaluation dataset and show that our model outperforms a strong extractive baseline.
- 124, TITLE: Weak Detection of Signal in the Spiked Wigner Model  
<http://proceedings.mlr.press/v97/chung19a.html>  
AUTHORS: Hye Won Chung, Ji Oon Lee  
HIGHLIGHT: In case the signal-to-noise ratio is under the threshold below which a reliable detection is impossible, we propose a hypothesis test based on the linear spectral statistics of the data matrix.
- 125, TITLE: New results on information theoretic clustering  
<http://proceedings.mlr.press/v97/cicalese19a.html>  
AUTHORS: Ferdinando Cicalese, Eduardo Laber, Lucas Murtinho  
HIGHLIGHT: We study the problem of optimizing the clustering of a set of vectors when the quality of the clustering is measured by the Entropy or the Gini impurity measure.
- 126, TITLE: Sensitivity Analysis of Linear Structural Causal Models  
<http://proceedings.mlr.press/v97/cinelli19a.html>  
AUTHORS: Carlos Cinelli, Daniel Kumor, Bryant Chen, Judea Pearl, Elias Bareinboim  
HIGHLIGHT: In this paper, we develop a formal, systematic approach to sensitivity analysis for arbitrary linear Structural Causal Models (SCMs).

- 127, TITLE: Dimensionality Reduction for Tukey Regression  
<http://proceedings.mlr.press/v97/clarkson19a.html>  
AUTHORS: Kenneth Clarkson, Ruosong Wang, David Woodruff  
HIGHLIGHT: We give the first dimensionality reduction methods for the overconstrained Tukey regression problem.
- 128, TITLE: On Medians of (Randomized) Pairwise Means  
<http://proceedings.mlr.press/v97/clemencon19a.html>  
AUTHORS: Stephan Clemencon, Pierre Laforgue, Patrice Bertail  
HIGHLIGHT: It is the purpose of this paper to extend this approach, in order to address other learning problems in particular, for which the performance criterion takes the form of an expectation over pairs of observations rather than over one single observation, as may be the case in pairwise ranking, clustering or metric learning.
- 129, TITLE: Quantifying Generalization in Reinforcement Learning  
<http://proceedings.mlr.press/v97/cobbe19a.html>  
AUTHORS: Karl Cobbe, Oleg Klimov, Chris Hesse, Taehoon Kim, John Schulman  
HIGHLIGHT: In this paper, we investigate the problem of overfitting in deep reinforcement learning.
- 130, TITLE: Empirical Analysis of Beam Search Performance Degradation in Neural Sequence Models  
<http://proceedings.mlr.press/v97/cohen19a.html>  
AUTHORS: Eldan Cohen, Christopher Beck  
HIGHLIGHT: We perform an empirical study of the behavior of beam search across three sequence synthesis tasks.
- 131, TITLE: Learning Linear-Quadratic Regulators Efficiently with only  $\sqrt{T}$  Regret  
<http://proceedings.mlr.press/v97/cohen19b.html>  
AUTHORS: Alon Cohen, Tomer Koren, Yishay Mansour  
HIGHLIGHT: We present the first computationally-efficient algorithm with  $\tilde{O}(\sqrt{T})$  regret for learning in Linear Quadratic Control systems with unknown dynamics.
- 132, TITLE: Certified Adversarial Robustness via Randomized Smoothing  
<http://proceedings.mlr.press/v97/cohen19c.html>  
AUTHORS: Jeremy Cohen, Elan Rosenfeld, Zico Kolter  
HIGHLIGHT: We show how to turn any classifier that classifies well under Gaussian noise into a new classifier that is certifiably robust to adversarial perturbations under the L2 norm.
- 133, TITLE: Gauge Equivariant Convolutional Networks and the Icosahedral CNN  
<http://proceedings.mlr.press/v97/cohen19d.html>  
AUTHORS: Taco Cohen, Maurice Weiler, Berkay Kicanaoglu, Max Welling  
HIGHLIGHT: Here we show how this principle can be extended beyond global symmetries to local gauge transformations.
- 134, TITLE: CURIOUS: Intrinsically Motivated Modular Multi-Goal Reinforcement Learning  
<http://proceedings.mlr.press/v97/colas19a.html>  
AUTHORS: Cédric Colas, Pierre-Yves Oudayer, Olivier Sigaud, Pierre Fournier, Mohamed Chetouani  
HIGHLIGHT: This paper proposes CURIOUS, an algorithm that leverages 1) a modular Universal Value Function Approximator with hindsight learning to achieve a diversity of goals of different kinds within a unique policy and 2) an automated curriculum learning mechanism that biases the attention of the agent towards goals maximizing the absolute learning progress.
- 135, TITLE: A fully differentiable beam search decoder  
<http://proceedings.mlr.press/v97/collobert19a.html>  
AUTHORS: Ronan Collobert, Awni Hannun, Gabriel Synnaeve



**HIGHLIGHT:** We introduce a new beam search decoder that is fully differentiable, making it possible to optimize at training time through the inference procedure.

136, **TITLE:** Scalable Metropolis-Hastings for Exact Bayesian Inference with Large Datasets  
<http://proceedings.mlr.press/v97/cornish19a.html>  
**AUTHORS:** Rob Cornish, Paul Vanetti, Alexandre Bouchard-Cote, George Deligiannidis, Arnaud Doucet  
**HIGHLIGHT:** We propose the Scalable Metropolis-Hastings (SMH) kernel that only requires processing on average  $O(1)$  or even  $O(1/\sqrt{n})$  data points per step.

137, **TITLE:** Adjustment Criteria for Generalizing Experimental Findings  
<http://proceedings.mlr.press/v97/correa19a.html>  
**AUTHORS:** Juan Correa, Jin Tian, Elias Bareinboim  
**HIGHLIGHT:** In this paper, we investigate the assumptions and machinery necessary for using covariate adjustment to correct for the biases generated by both of these problems, and generalize experimental data to infer causal effects in a new domain.

138, **TITLE:** Online Learning with Sleeping Experts and Feedback Graphs  
<http://proceedings.mlr.press/v97/cortes19a.html>  
**AUTHORS:** Corinna Cortes, Giulia Desalvo, Claudio Gentile, Mehryar Mohri, Scott Yang  
**HIGHLIGHT:** Our main contribution is then to relax this assumption, present a more general notion of sleeping regret, and derive a general algorithm with strong theoretical guarantees.

139, **TITLE:** Active Learning with Disagreement Graphs  
<http://proceedings.mlr.press/v97/cortes19b.html>  
**AUTHORS:** Corinna Cortes, Giulia Desalvo, Mehryar Mohri, Ningshan Zhang, Claudio Gentile  
**HIGHLIGHT:** We present two novel enhancements of an online importance-weighted active learning algorithm IWAL, using the properties of disagreements among hypotheses.

140, **TITLE:** Shape Constraints for Set Functions  
<http://proceedings.mlr.press/v97/cotter19a.html>  
**AUTHORS:** Andrew Cotter, Maya Gupta, Heinrich Jiang, Erez Louidor, James Muller, Tamann Narayan, Serena Wang, Tao Zhu  
**HIGHLIGHT:** We propose making set functions more understandable and regularized by capturing domain knowledge through shape constraints.

141, **TITLE:** Training Well-Generalizing Classifiers for Fairness Metrics and Other Data-Dependent Constraints  
<http://proceedings.mlr.press/v97/cotter19b.html>  
**AUTHORS:** Andrew Cotter, Maya Gupta, Heinrich Jiang, Nathan Srebro, Karthik Sridharan, Serena Wang, Blake Woodworth, Seungil You  
**HIGHLIGHT:** To improve generalization, we frame the problem as a two-player game where one player optimizes the model parameters on a training dataset, and the other player enforces the constraints on an independent validation dataset.

142, **TITLE:** Monge blunts Bayes: Hardness Results for Adversarial Training  
<http://proceedings.mlr.press/v97/cranko19a.html>  
**AUTHORS:** Zac Cranko, Aditya Menon, Richard Nock, Cheng Soon Ong, Zhan Shi, Christian Walder  
**HIGHLIGHT:** We suggest a formal answer for losses that satisfy the minimal statistical requirement of being proper.

143, **TITLE:** Boosted Density Estimation Remastered  
<http://proceedings.mlr.press/v97/cranko19b.html>  
**AUTHORS:** Zac Cranko, Richard Nock

**HIGHLIGHT:** We show how to combine this latter approach and the classical boosting theory in supervised learning to get the first density estimation algorithm that provably achieves geometric convergence under very weak assumptions.

144, **TITLE:** Submodular Cost Submodular Cover with an Approximate Oracle  
<http://proceedings.mlr.press/v97/crawford19a.html>

**AUTHORS:** Victoria Crawford, Alan Kuhnle, My Thai

**HIGHLIGHT:** In this work, we study the Submodular Cost Submodular Cover problem, which is to minimize the submodular cost required to ensure that the submodular benefit function exceeds a given threshold.

145, **TITLE:** Flexibly Fair Representation Learning by Disentanglement  
<http://proceedings.mlr.press/v97/creager19a.html>

**AUTHORS:** Elliot Creager, David Madras, Joern-Henrik Jacobsen, Marissa Weis, Kevin Swersky, Toniann Pitassi, Richard Zemel

**HIGHLIGHT:** Taking inspiration from the disentangled representation learning literature, we propose an algorithm for learning compact representations of datasets that are useful for reconstruction and prediction, but are also flexibly fair, meaning they can be easily modified at test time to achieve subgroup demographic parity with respect to multiple sensitive attributes and their conjunctions.

146, **TITLE:** Anytime Online-to-Batch, Optimism and Acceleration

<http://proceedings.mlr.press/v97/cutkosky19a.html>

**AUTHORS:** Ashok Cutkosky

**HIGHLIGHT:** We close this gap by introducing a black-box modification to any online learning algorithm whose iterates converge to the optimum in stochastic scenarios.

147, **TITLE:** Matrix-Free Preconditioning in Online Learning

<http://proceedings.mlr.press/v97/cutkosky19b.html>

**AUTHORS:** Ashok Cutkosky, Tamas Sarlos

**HIGHLIGHT:** We provide an online convex optimization algorithm with regret that interpolates between the regret of an algorithm using an optimal preconditioning matrix and one using a diagonal preconditioning matrix.

148, **TITLE:** Minimal Achievable Sufficient Statistic Learning

<http://proceedings.mlr.press/v97/cvitkovic19a.html>

**AUTHORS:** Milan Cvitkovic, G?nther Koliander

**HIGHLIGHT:** We introduce Minimal Achievable Sufficient Statistic (MASS) Learning, a machine learning training objective for which the minima are minimal sufficient statistics with respect to a class of functions being optimized over (e.g., deep networks).

149, **TITLE:** Open Vocabulary Learning on Source Code with a Graph-Structured Cache

<http://proceedings.mlr.press/v97/cvitkovic19b.html>

**AUTHORS:** Milan Cvitkovic, Badal Singh, Animashree Anandkumar

**HIGHLIGHT:** We introduce a Graph-Structured Cache to address this problem; this cache contains a node for each new word the model encounters with edges connecting each word to its occurrences in the code.

150, **TITLE:** The Value Function Polytope in Reinforcement Learning

<http://proceedings.mlr.press/v97/dadashi19a.html>

**AUTHORS:** Robert Dadashi, Marc G. Bellemare, Adrien Ali Taiga, Nicolas Le Roux, Dale Schuurmans

**HIGHLIGHT:** Our main contribution is the characterization of the nature of its shape: a general polytope (Aigner et al., 2010).

151, **TITLE:** Bayesian Optimization Meets Bayesian Optimal Stopping

<http://proceedings.mlr.press/v97/dai19a.html>

**AUTHORS:** Zhongxiang Dai, Haibin Yu, Bryan Kian Hsiang Low, Patrick Jaillet

**HIGHLIGHT:** This paper proposes to unify BO (specifically, Gaussian process-upper confidence bound (GP-UCB)) with Bayesian optimal stopping (BO-BOS) to boost the epoch efficiency of BO.

152, **TITLE:** Policy Certificates: Towards Accountable Reinforcement Learning  
<http://proceedings.mlr.press/v97/dann19a.html>  
**AUTHORS:** Christoph Dann, Lihong Li, Wei Wei, Emma Brunskill  
**HIGHLIGHT:** We address this lack of accountability by proposing that algorithms output policy certificates.

153, **TITLE:** Learning Fast Algorithms for Linear Transforms Using Butterfly Factorizations  
<http://proceedings.mlr.press/v97/dao19a.html>  
**AUTHORS:** Tri Dao, Albert Gu, Matthew Eichhorn, Atri Rudra, Christopher Re  
**HIGHLIGHT:** Motivated by a characterization of fast matrix-vector multiplication as products of sparse matrices, we introduce a parameterization of divide-and-conquer methods that is capable of representing a large class of transforms.

154, **TITLE:** A Kernel Theory of Modern Data Augmentation  
<http://proceedings.mlr.press/v97/dao19b.html>  
**AUTHORS:** Tri Dao, Albert Gu, Alexander Ratner, Virginia Smith, Chris De Sa, Christopher Re  
**HIGHLIGHT:** In this paper, we seek to establish a theoretical framework for understanding data augmentation.

155, **TITLE:** TarMAC: Targeted Multi-Agent Communication  
<http://proceedings.mlr.press/v97/das19a.html>  
**AUTHORS:** Abhishek Das, Th?ophile Gervet, Joshua Romoff, Dhruv Batra, Devi Parikh, Mike Rabbat, Joelle Pineau  
**HIGHLIGHT:** We propose a targeted communication architecture for multi-agent reinforcement learning, where agents learn both what messages to send and whom to address them to while performing cooperative tasks in partially-observable environments.

156, **TITLE:** Teaching a black-box learner  
<http://proceedings.mlr.press/v97/dasgupta19a.html>  
**AUTHORS:** Sanjoy Dasgupta, Daniel Hsu, Stefanos Poulis, Xiaojin Zhu  
**HIGHLIGHT:** We consider the problem of teaching a learner whose representation and hypothesis class are unknown—that is, the learner is a black box.

157, **TITLE:** Stochastic Deep Networks  
<http://proceedings.mlr.press/v97/de-bie19a.html>  
**AUTHORS:** Gwendoline De Bie, Gabriel Peyr?, Marco Cuturi  
**HIGHLIGHT:** We propose in this work a deep framework designed to handle crucial aspects of measures, namely permutation invariances, variations in weights and cardinality.

158, **TITLE:** Learning-to-Learn Stochastic Gradient Descent with Biased Regularization  
<http://proceedings.mlr.press/v97/denevi19a.html>  
**AUTHORS:** Giulia Denevi, Carlo Ciliberto, Riccardo Grazi, Massimiliano Pontil  
**HIGHLIGHT:** We present an average excess risk bound for such a learning algorithm that quantifies the potential benefit of using a bias vector with respect to the unbiased case.

159, **TITLE:** A Multitask Multiple Kernel Learning Algorithm for Survival Analysis with Application to Cancer Biology  
<http://proceedings.mlr.press/v97/dereli19a.html>  
**AUTHORS:** Onur Dereli, Ceyda Oguz, Mehmet G?nen  
**HIGHLIGHT:** Rather than performing survival analysis on each data set to predict survival times of cancer patients, we developed a novel multitask approach based on multiple kernel learning (MKL).

- 160, TITLE: Learning to Convolve: A Generalized Weight-Tying Approach  
<http://proceedings.mlr.press/v97/diaconu19a.html>  
AUTHORS: Nichita Diaconu, Daniel Worrall  
HIGHLIGHT: In this paper, we learn how to transform filters for use in the group convolution, focussing on rotation.
- 161, TITLE: Sever: A Robust Meta-Algorithm for Stochastic Optimization  
<http://proceedings.mlr.press/v97/diakonikolas19a.html>  
AUTHORS: Ilias Diakonikolas, Gautam Kamath, Daniel Kane, Jerry Li, Jacob Steinhardt, Alistair Stewart  
HIGHLIGHT: To address this, we introduce a new meta-algorithm that can take in a base learner such as least squares or stochastic gradient descent, and harden the learner to be resistant to outliers.
- 162, TITLE: Approximated Oracle Filter Pruning for Destructive CNN Width Optimization  
<http://proceedings.mlr.press/v97/ding19a.html>  
AUTHORS: Xiaohan Ding, Guiguang Ding, Yuchen Guo, Jungong Han, Chenggang Yan  
HIGHLIGHT: To address these problems, we propose Approximated Oracle Filter Pruning (AOFPP), which keeps searching for the least important filters in a binary search manner, makes pruning attempts by masking out filters randomly, accumulates the resulting errors, and finetunes the model via a multi-path framework.
- 163, TITLE: Noisy Dual Principal Component Pursuit  
<http://proceedings.mlr.press/v97/ding19b.html>  
AUTHORS: Tianyu Ding, Zihui Zhu, Tianjiao Ding, Yunchen Yang, Daniel Robinson, Manolis Tsakiris, Rene Vidal  
HIGHLIGHT: Noisy Dual Principal Component Pursuit
- 164, TITLE: Finite-Time Analysis of Distributed TD(0) with Linear Function Approximation on Multi-Agent Reinforcement Learning  
<http://proceedings.mlr.press/v97/doan19a.html>  
AUTHORS: Thinh Doan, Siva Maguluri, Justin Romberg  
HIGHLIGHT: Our main contribution is providing a finite-time analysis for the convergence of the distributed TD(0) algorithm.
- 165, TITLE: Trajectory-Based Off-Policy Deep Reinforcement Learning  
<http://proceedings.mlr.press/v97/doerr19a.html>  
AUTHORS: Andreas Doerr, Michael Volpp, Marc Toussaint, Trimpe Sebastian, Christian Daniel  
HIGHLIGHT: This work addresses these weaknesses by combining recent improvements in the reuse of off-policy data and exploration in parameter space with deterministic behavioral policies.
- 166, TITLE: Generalized No Free Lunch Theorem for Adversarial Robustness  
<http://proceedings.mlr.press/v97/dohmatob19a.html>  
AUTHORS: Elvis Dohmatob  
HIGHLIGHT: This manuscript presents some new impossibility results on adversarial robustness in machine learning, a very important yet largely open problem.
- 167, TITLE: Width Provably Matters in Optimization for Deep Linear Neural Networks  
<http://proceedings.mlr.press/v97/du19a.html>  
AUTHORS: Simon Du, Wei Hu  
HIGHLIGHT: We prove that for an  $L$ -layer fully-connected linear neural network, if the width of every hidden layer is  $\widetilde{\Omega}(L \cdot r \cdot d_{\text{out}} \cdot \kappa^3)$ , where  $r$  and  $\kappa$  are the rank and the condition number of the input data, and  $d_{\text{out}}$  is the output dimension, then gradient descent with Gaussian random initialization converges to a global minimum at a linear rate.
- 168, TITLE: Provably efficient RL with Rich Observations via Latent State Decoding  
<http://proceedings.mlr.press/v97/du19b.html>

AUTHORS: Simon Du, Akshay Krishnamurthy, Nan Jiang, Alekh Agarwal, Miroslav Dudik, John Langford  
HIGHLIGHT: We study the exploration problem in episodic MDPs with rich observations generated from a small number of latent states.

169, TITLE: Gradient Descent Finds Global Minima of Deep Neural Networks  
<http://proceedings.mlr.press/v97/du19c.html>  
AUTHORS: Simon Du, Jason Lee, Haochuan Li, Liwei Wang, Xiyu Zhai  
HIGHLIGHT: The current paper proves gradient descent achieves zero training loss in polynomial time for a deep over-parameterized neural network with residual connections (ResNet).

170, TITLE: Incorporating Grouping Information into Bayesian Decision Tree Ensembles  
<http://proceedings.mlr.press/v97/du19d.html>  
AUTHORS: Junliang Du, Antonio Linero  
HIGHLIGHT: We consider the problem of nonparametric regression in the high-dimensional setting in which  $\mathbb{P}$  is  $\epsilon$ -N $\mathbb{S}$ .

171, TITLE: Task-Agnostic Dynamics Priors for Deep Reinforcement Learning  
<http://proceedings.mlr.press/v97/du19e.html>  
AUTHORS: Yilun Du, Karthic Narasimhan  
HIGHLIGHT: In this work, we propose an approach to learn task-agnostic dynamics priors from videos and incorporate them into an RL agent.

172, TITLE: Optimal Auctions through Deep Learning  
<http://proceedings.mlr.press/v97/duetting19a.html>  
AUTHORS: Paul Duetting, Zhe Feng, Harikrishna Narasimhan, David Parkes, Sai Srivatsa Ravindranath  
HIGHLIGHT: In this work, we initiate the exploration of the use of tools from deep learning for the automated design of optimal auctions.

173, TITLE: Wasserstein of Wasserstein Loss for Learning Generative Models  
<http://proceedings.mlr.press/v97/dukler19a.html>  
AUTHORS: Yonatan Dukler, Wuchen Li, Alex Lin, Guido Montufar  
HIGHLIGHT: We propose to use the Wasserstein distance itself as the ground metric on the sample space of images.

174, TITLE: Learning interpretable continuous-time models of latent stochastic dynamical systems  
<http://proceedings.mlr.press/v97/duncker19a.html>  
AUTHORS: Lea Duncker, Gergo Bohner, Julien Boussard, Maneesh Sahani  
HIGHLIGHT: This form yields a flexible nonparametric model of the dynamics, with a representation corresponding directly to the interpretable portraits routinely employed in the study of nonlinear dynamical systems.

175, TITLE: Autoregressive Energy Machines  
<http://proceedings.mlr.press/v97/durkan19a.html>  
AUTHORS: Conor Durkan, Charlie Nash  
HIGHLIGHT: We propose the Autoregressive Energy Machine, an energy-based model which simultaneously learns an unnormalized density and computes an importance-sampling estimate of the normalizing constant for each conditional in an autoregressive decomposition.

176, TITLE: Band-limited Training and Inference for Convolutional Neural Networks  
<http://proceedings.mlr.press/v97/dziedzic19a.html>  
AUTHORS: Adam Dziedzic, John Paparrizos, Sanjay Krishnan, Aaron Elmore, Michael Franklin  
HIGHLIGHT: We explore artificially constraining the frequency spectra of these filters and data, called band-limiting, during training.

- 177, TITLE: Imitating Latent Policies from Observation  
<http://proceedings.mlr.press/v97/edwards19a.html>  
AUTHORS: Ashley Edwards, Himanshu Sahni, Yannick Schroecker, Charles Isbell  
HIGHLIGHT: In this paper, we describe a novel approach to imitation learning that infers latent policies directly from state observations.
- 178, TITLE: Semi-Cyclic Stochastic Gradient Descent  
<http://proceedings.mlr.press/v97/eichner19a.html>  
AUTHORS: Hubert Eichner, Tomer Koren, Brendan McMahan, Nathan Srebro, Kunal Talwar  
HIGHLIGHT: We show that such block-cyclic structure can significantly deteriorate the performance of SGD, but propose a simple approach that allows prediction with the same guarantees as for i.i.d., non-cyclic, sampling.
- 179, TITLE: GDPP: Learning Diverse Generations using Determinantal Point Processes  
<http://proceedings.mlr.press/v97/elfeki19a.html>  
AUTHORS: Mohamed Elfeki, Camille Couprie, Morgane Riviere, Mohamed Elhoseiny  
HIGHLIGHT: In this work, we draw inspiration from Determinantal Point Process (DPP) to propose an unsupervised penalty loss that alleviates mode collapse while producing higher quality samples.
- 180, TITLE: Sequential Facility Location: Approximate Submodularity and Greedy Algorithm  
<http://proceedings.mlr.press/v97/elhamifar19a.html>  
AUTHORS: Ehsan Elhamifar  
HIGHLIGHT: We propose a cardinality-constrained sequential facility location function that finds a fixed number of representatives, where the sequence of representatives is compatible with the dynamic model and well encodes the data.
- 181, TITLE: Improved Convergence for  $\ell_1$  and  $\ell_8$  Regression via Iteratively Reweighted Least Squares  
<http://proceedings.mlr.press/v97/ene19a.html>  
AUTHORS: Alina Ene, Adrian Vladu  
HIGHLIGHT: In this paper we propose a simple and natural version of IRLS for solving  $\ell_\infty$  and  $\ell_1$  regression, which provably converges to a  $(1+\epsilon)$ -approximate solution in  $O(m^{1/3} \log(1/\epsilon) \epsilon^{2/3} + \log m \epsilon^2)$  iterations, where  $m$  is the number of rows of the input matrix.
- 182, TITLE: Exploring the Landscape of Spatial Robustness  
<http://proceedings.mlr.press/v97/engstrom19a.html>  
AUTHORS: Logan Engstrom, Brandon Tran, Dimitris Tsipras, Ludwig Schmidt, Aleksander Madry  
HIGHLIGHT: In this work, we thoroughly investigate the vulnerability of neural network-based classifiers to rotations and translations.
- 183, TITLE: Cross-Domain 3D Equivariant Image Embeddings  
<http://proceedings.mlr.press/v97/esteves19a.html>  
AUTHORS: Carlos Esteves, Avneesh Sud, Zhengyi Luo, Kostas Daniilidis, Ameet Makadia  
HIGHLIGHT: In this paper we learn 2D image embeddings with a similar equivariant structure: embedding the image of a 3D object should commute with rotations of the object.
- 184, TITLE: On the Connection Between Adversarial Robustness and Saliency Map Interpretability  
<http://proceedings.mlr.press/v97/etmann19a.html>  
AUTHORS: Christian Etmann, Sebastian Lunz, Peter Maass, Carola Schoenlieb  
HIGHLIGHT: We aim to quantify this behaviour by considering the alignment between input image and saliency map.
- 185, TITLE: Non-monotone Submodular Maximization with Nearly Optimal Adaptivity and Query Complexity  
<http://proceedings.mlr.press/v97/fahrback19a.html>

AUTHORS: Matthew Fahrback, Vahab Mirrokni, Morteza Zadimoghaddam  
HIGHLIGHT: In this paper, we give the first constant-factor approximation algorithm for maximizing a non-monotone submodular function subject to a cardinality constraint  $k$  that runs in  $O(\log(n))$  adaptive rounds and makes  $O(n \log(k))$  oracle queries in expectation.

186, TITLE: Multi-Frequency Vector Diffusion Maps  
<http://proceedings.mlr.press/v97/fan19a.html>  
AUTHORS: Yifeng Fan, Zhizhen Zhao  
HIGHLIGHT: We introduce multi-frequency vector diffusion maps (MFVDM), a new framework for organizing and analyzing high dimensional data sets.

187, TITLE: Stable-Predictive Optimistic Counterfactual Regret Minimization  
<http://proceedings.mlr.press/v97/farina19a.html>  
AUTHORS: Gabriele Farina, Christian Kroer, Noam Brown, Tuomas Sandholm  
HIGHLIGHT: In this work we present the first CFR variant that breaks the square-root dependence on iterations.

188, TITLE: Regret Circuits: Composability of Regret Minimizers  
<http://proceedings.mlr.press/v97/farina19b.html>  
AUTHORS: Gabriele Farina, Christian Kroer, Tuomas Sandholm  
HIGHLIGHT: In this paper we study the general composability of regret minimizers.

189, TITLE: Dead-ends and Secure Exploration in Reinforcement Learning  
<http://proceedings.mlr.press/v97/fatemi19a.html>  
AUTHORS: Mehdi Fatemi, Shikhar Sharma, Harm Van Seijen, Samira Ebrahimi Kahou  
HIGHLIGHT: To deal with the bridge effect, we propose a condition for exploration, called security.

190, TITLE: Invariant-Equivariant Representation Learning for Multi-Class Data  
<http://proceedings.mlr.press/v97/feige19a.html>  
AUTHORS: Ilya Feige  
HIGHLIGHT: We introduce an approach to probabilistic modelling that learns to represent data with two separate deep representations: an invariant representation that encodes the information of the class from which the data belongs, and an equivariant representation that encodes the symmetry transformation defining the particular data point within the class manifold (equivariant in the sense that the representation varies naturally with symmetry transformations).

191, TITLE: The advantages of multiple classes for reducing overfitting from test set reuse  
<http://proceedings.mlr.press/v97/feldman19a.html>  
AUTHORS: Vitaly Feldman, Roy Frostig, Moritz Hardt  
HIGHLIGHT: We show a new upper bound of  $O(\max\{\sqrt{k \log(n)/(mn)}, k/n\})$  on the worst-case bias that any attack can achieve in a prediction problem with  $m$  classes.

192, TITLE: Decentralized Exploration in Multi-Armed Bandits  
<http://proceedings.mlr.press/v97/feraud19a.html>  
AUTHORS: Raphael Feraud, Reda Alami, Romain Laroche  
HIGHLIGHT: We consider the decentralized exploration problem: a set of players collaborate to identify the best arm by asynchronously interacting with the same stochastic environment.

193, TITLE: Almost surely constrained convex optimization  
<http://proceedings.mlr.press/v97/fercoq19a.html>  
AUTHORS: Olivier Fercoq, Ahmet Alacaoglu, Ion Necoara, Volkan Cevher  
HIGHLIGHT: We propose a stochastic gradient framework for solving stochastic composite convex optimization problems with (possibly) infinite number of linear inclusion constraints that need to be satisfied almost surely.

194, TITLE: Online Meta-Learning

<http://proceedings.mlr.press/v97/finn19a.html>

AUTHORS: Chelsea Finn, Aravind Rajeswaran, Sham Kakade, Sergey Levine

HIGHLIGHT: This work introduces an online meta-learning setting, which merges ideas from both paradigms to better capture the spirit and practice of continual lifelong learning.

195, TITLE: DL2: Training and Querying Neural Networks with Logic

<http://proceedings.mlr.press/v97/fischer19a.html>

AUTHORS: Marc Fischer, Mislav Balunovic, Dana Drachler-Cohen, Timon Gehr, Ce Zhang, Martin Vechev

HIGHLIGHT: We present DL2, a system for training and querying neural networks with logical constraints.

196, TITLE: Bayesian Action Decoder for Deep Multi-Agent Reinforcement Learning

<http://proceedings.mlr.press/v97/foerster19a.html>

AUTHORS: Jakob Foerster, Francis Song, Edward Hughes, Neil Burch, Iain Dunning, Shimon Whiteson, Matthew Botvinick, Michael Bowling

HIGHLIGHT: We present the Bayesian action decoder (BAD), a new multi-agent learning method that uses an approximate Bayesian update to obtain a public belief that conditions on the actions taken by all agents in the environment.

197, TITLE: Scalable Nonparametric Sampling from Multimodal Posteriors with the Posterior Bootstrap

<http://proceedings.mlr.press/v97/fong19a.html>

AUTHORS: Edwin Fong, Simon Lyddon, Chris Holmes

HIGHLIGHT: We present a scalable Bayesian nonparametric learning routine that enables posterior sampling through the optimization of suitably randomized objective functions.

198, TITLE: On discriminative learning of prediction uncertainty

<http://proceedings.mlr.press/v97/franc19a.html>

AUTHORS: Vojtech Franc, Daniel Prusa

HIGHLIGHT: We propose a discriminative algorithm learning an uncertainty function which preserves ordering of the input space induced by the conditional risk, and hence can be used to construct optimal rejection strategies.

199, TITLE: Learning Discrete Structures for Graph Neural Networks

<http://proceedings.mlr.press/v97/franceschi19a.html>

AUTHORS: Luca Franceschi, Mathias Niepert, Massimiliano Pontil, Xiao He

HIGHLIGHT: With this work, we propose to jointly learn the graph structure and the parameters of graph convolutional networks (GCNs) by approximately solving a bilevel program that learns a discrete probability distribution on the edges of the graph.

200, TITLE: Distributional Multivariate Policy Evaluation and Exploration with the Bellman GAN

<http://proceedings.mlr.press/v97/freirich19a.html>

AUTHORS: Dror Freirich, Tzahi Shimkin, Ron Meir, Aviv Tamar

HIGHLIGHT: In this work, we show that the distributional Bellman equation, which drives DiRL methods, is equivalent to a generative adversarial network (GAN) model.

201, TITLE: Approximating Orthogonal Matrices with Effective Givens Factorization

<http://proceedings.mlr.press/v97/frerix19a.html>

AUTHORS: Thomas Frerix, Joan Bruna

HIGHLIGHT: We analyze effective approximation of unitary matrices.

202, TITLE: Fast and Flexible Inference of Joint Distributions from their Marginals

<http://proceedings.mlr.press/v97/frogner19a.html>

AUTHORS: Charlie Frogner, Tomaso Poggio

HIGHLIGHT: In this paper, we treat the inference problem generally and propose a unified class of models that encompasses some of those previously proposed while including many new ones.



203, TITLE: Analyzing and Improving Representations with the Soft Nearest Neighbor Loss  
<http://proceedings.mlr.press/v97/frosst19a.html>  
AUTHORS: Nicholas Frosst, Nicolas Papernot, Geoffrey Hinton  
HIGHLIGHT: We explore and expand the Soft Nearest Neighbor Loss to measure the entanglement of class manifolds in representation space: i.e., how close pairs of points from the same class are relative to pairs of points from different classes.

204, TITLE: Diagnosing Bottlenecks in Deep Q-learning Algorithms  
<http://proceedings.mlr.press/v97/fu19a.html>  
AUTHORS: Justin Fu, Aviral Kumar, Matthew Soh, Sergey Levine  
HIGHLIGHT: In this work, we aim to experimentally investigate potential issues in Q-learning, by means of a "unit testing" framework where we can utilize oracles to disentangle sources of error.

205, TITLE: MetricGAN: Generative Adversarial Networks based Black-box Metric Scores Optimization for Speech Enhancement  
<http://proceedings.mlr.press/v97/fu19b.html>  
AUTHORS: Szu-Wei Fu, Chien-Feng Liao, Yu Tsao, Shou-De Lin  
HIGHLIGHT: To overcome this issue, we propose a novel MetricGAN approach with an aim to optimize the generator with respect to one or multiple evaluation metrics.

206, TITLE: Beyond Adaptive Submodularity: Approximation Guarantees of Greedy Policy with Adaptive Submodularity Ratio  
<http://proceedings.mlr.press/v97/fujii19a.html>  
AUTHORS: Kaito Fujii, Shinsaku Sakae  
HIGHLIGHT: We propose a new concept named adaptive submodularity ratio to study the greedy policy for sequential decision making.

207, TITLE: Off-Policy Deep Reinforcement Learning without Exploration  
<http://proceedings.mlr.press/v97/fujimoto19a.html>  
AUTHORS: Scott Fujimoto, David Meger, Doina Precup  
HIGHLIGHT: In this paper, we demonstrate that due to errors introduced by extrapolation, standard off-policy deep reinforcement learning algorithms, such as DQN and DDPG, are incapable of learning with data uncorrelated to the distribution under the current policy, making them ineffective for this fixed batch setting.

208, TITLE: Transfer Learning for Related Reinforcement Learning Tasks via Image-to-Image Translation  
<http://proceedings.mlr.press/v97/gamrian19a.html>  
AUTHORS: Shani Gamrian, Yoav Goldberg  
HIGHLIGHT: We demonstrate the approach on synthetic visual variants of the Breakout game, as well as on transfer between subsequent levels of Road Fighter, a Nintendo car-driving game.

209, TITLE: Breaking the Softmax Bottleneck via Learnable Monotonic Pointwise Non-linearities  
<http://proceedings.mlr.press/v97/ganea19a.html>  
AUTHORS: Octavian Ganea, Sylvain Gelly, Gary Becigneul, Aliaksei Severyn  
HIGHLIGHT: As an efficient and effective solution to alleviate this issue, we propose to learn parametric monotonic functions on top of the logits.

210, TITLE: Graph U-Nets  
<http://proceedings.mlr.press/v97/gao19a.html>  
AUTHORS: Hongyang Gao, Shuiwang Ji  
HIGHLIGHT: To address these challenges, we propose novel graph pooling (gPool) and unpooling (gUnpool) operations in this work.

211, TITLE: Deep Generative Learning via Variational Gradient Flow

<http://proceedings.mlr.press/v97/gao19b.html>

AUTHORS: Yuan Gao, Yuling Jiao, Yang Wang, Yao Wang, Can Yang, Shunkang Zhang

HIGHLIGHT: We propose a framework to learn deep generative models via variational Gradient Flows (VGF) on probability spaces.

212, TITLE: Rate Distortion For Model Compression: From Theory To Practice

<http://proceedings.mlr.press/v97/gao19c.html>

AUTHORS: Weihao Gao, Yu-Han Liu, Chong Wang, Sewoong Oh

HIGHLIGHT: In this paper, we propose principled approaches to improve upon the common heuristics used in those building blocks, by studying the fundamental limit for model compression via the rate distortion theory.

213, TITLE: Demystifying Dropout

<http://proceedings.mlr.press/v97/gao19d.html>

AUTHORS: Hongchang Gao, Jian Pei, Heng Huang

HIGHLIGHT: In this paper, unlike existing works, we explore it from a new perspective to provide new insight into this line of research.

214, TITLE: Geometric Scattering for Graph Data Analysis

<http://proceedings.mlr.press/v97/gao19e.html>

AUTHORS: Feng Gao, Guy Wolf, Matthew Hirn

HIGHLIGHT: We explore the generalization of scattering transforms from traditional (e.g., image or audio) signals to graph data, analogous to the generalization of ConvNets in geometric deep learning, and the utility of extracted graph features in graph data analysis.

215, TITLE: Multi-Frequency Phase Synchronization

<http://proceedings.mlr.press/v97/gao19f.html>

AUTHORS: Tingran Gao, Zhizhen Zhao

HIGHLIGHT: We propose a novel formulation for phase synchronization—the statistical problem of jointly estimating alignment angles from noisy pairwise comparisons—as a nonconvex optimization problem that enforces consistency among the pairwise comparisons in multiple frequency channels.

216, TITLE: Optimal Mini-Batch and Step Sizes for SAGA

<http://proceedings.mlr.press/v97/gazagnadou19a.html>

AUTHORS: Nidham Gazagnadou, Robert Gower, Joseph Salmon

HIGHLIGHT: Using these bounds, and since the SAGA algorithm is part of this JacSketch family, we suggest a new standard practice for setting the step and mini-batch sizes for SAGA that are competitive with a numerical grid search.

217, TITLE: SelectiveNet: A Deep Neural Network with an Integrated Reject Option

<http://proceedings.mlr.press/v97/geifman19a.html>

AUTHORS: Yonatan Geifman, Ran El-Yaniv

HIGHLIGHT: We consider the problem of selective prediction (also known as reject option) in deep neural networks, and introduce SelectiveNet, a deep neural architecture with an integrated reject option.

218, TITLE: A Theory of Regularized Markov Decision Processes

<http://proceedings.mlr.press/v97/geist19a.html>

AUTHORS: Matthieu Geist, Bruno Scherrer, Olivier Pietquin

HIGHLIGHT: We propose a general theory of regularized Markov Decision Processes that generalizes these approaches in two directions: we consider a larger class of regularizers, and we consider the general modified policy iteration approach, encompassing both policy iteration and value iteration.

219, TITLE: DeepMDP: Learning Continuous Latent Space Models for Representation Learning

<http://proceedings.mlr.press/v97/gelada19a.html>

AUTHORS: Carles Gelada, Saurabh Kumar, Jacob Buckman, Ofir Nachum, Marc G. Bellemare

**HIGHLIGHT:** To formalize this process, we introduce the concept of a  $\text{DeepMDP}$ , a parameterized latent space model that is trained via the minimization of two tractable latent space losses: prediction of rewards and prediction of the distribution over next latent states.

220, **TITLE:** Partially Linear Additive Gaussian Graphical Models  
<http://proceedings.mlr.press/v97/geng19a.html>  
**AUTHORS:** Sinong Geng, Minhao Yan, Mladen Kolar, Sanmi Koyejo  
**HIGHLIGHT:** We propose a partially linear additive Gaussian graphical model (PLA-GGM) for the estimation of associations between random variables distorted by observed confounders.

221, **TITLE:** Learning and Data Selection in Big Datasets  
<http://proceedings.mlr.press/v97/ghadikolaei19a.html>  
**AUTHORS:** Hossein Shokri Ghadikolaei, Hadi Ghauch, Carlo Fischione, Mikael Skoglund  
**HIGHLIGHT:** More specifically, we propose a framework that jointly learns the input-output mapping as well as the most representative samples of the dataset (sufficient dataset).

222, **TITLE:** Improved Parallel Algorithms for Density-Based Network Clustering  
<http://proceedings.mlr.press/v97/ghaffari19a.html>  
**AUTHORS:** Mohsen Ghaffari, Silvio Lattanzi, Slobodan Mitrovic  
**HIGHLIGHT:** In the case of  $k$ -core decomposition, our work improves exponentially on the algorithm provided by Esfandiari et al. (ICML'18).

223, **TITLE:** Recursive Sketches for Modular Deep Learning  
<http://proceedings.mlr.press/v97/ghazi19a.html>  
**AUTHORS:** Badih Ghazi, Rina Panigrahy, Joshua Wang  
**HIGHLIGHT:** We present a mechanism to compute a sketch (succinct summary) of how a complex modular deep network processes its inputs.

224, **TITLE:** An Instability in Variational Inference for Topic Models  
<http://proceedings.mlr.press/v97/ghorbani19a.html>  
**AUTHORS:** Behrooz Ghorbani, Hamid Javadi, Andrea Montanari  
**HIGHLIGHT:** We show that these methods suffer from an instability that can produce misleading conclusions.

225, **TITLE:** An Investigation into Neural Net Optimization via Hessian Eigenvalue Density  
<http://proceedings.mlr.press/v97/ghorbani19b.html>  
**AUTHORS:** Behrooz Ghorbani, Shankar Krishnan, Ying Xiao  
**HIGHLIGHT:** To understand the dynamics of training in deep neural networks, we study the evolution of the Hessian eigenvalue density throughout the optimization process.

226, **TITLE:** Data Shapley: Equitable Valuation of Data for Machine Learning  
<http://proceedings.mlr.press/v97/ghorbani19c.html>  
**AUTHORS:** Amirata Ghorbani, James Zou  
**HIGHLIGHT:** In this work, we develop a principled framework to address data valuation in the context of supervised machine learning.

227, **TITLE:** Efficient Dictionary Learning with Gradient Descent  
<http://proceedings.mlr.press/v97/gilboa19a.html>  
**AUTHORS:** Dar Gilboa, Sam Buchanan, John Wright  
**HIGHLIGHT:** We study one such problem – complete orthogonal dictionary learning, and provide convergence guarantees for randomly initialized gradient descent to the neighborhood of a global optimum.

228, **TITLE:** A Tree-Based Method for Fast Repeated Sampling of Determinantal Point Processes  
<http://proceedings.mlr.press/v97/gillenwater19a.html>

AUTHORS: Jennifer Gillenwater, Alex Kulesza, Zelda Mariet, Sergei Vassilvtiskii  
HIGHLIGHT: In this work we address both of these shortcomings.

229, TITLE: Learning to Groove with Inverse Sequence Transformations  
<http://proceedings.mlr.press/v97/gillick19a.html>  
AUTHORS: Jon Gillick, Adam Roberts, Jesse Engel, Douglas Eck, David Bamman  
HIGHLIGHT: We explore models for translating abstract musical ideas (scores, rhythms) into expressive performances using seq2seq and recurrent variational information bottleneck (VIB) models. Focusing on the case of drum set players, we create and release a new dataset for this purpose, containing over 13 hours of recordings by professional drummers aligned with fine-grained timing and dynamics information.

230, TITLE: Adversarial Examples Are a Natural Consequence of Test Error in Noise  
<http://proceedings.mlr.press/v97/gilmer19a.html>  
AUTHORS: Justin Gilmer, Nicolas Ford, Nicholas Carlini, Ekin Cubuk  
HIGHLIGHT: In this paper we provide both empirical and theoretical evidence that these are two manifestations of the same underlying phenomenon, and therefore the adversarial robustness and corruption robustness research programs are closely related.

231, TITLE: Discovering Conditionally Salient Features with Statistical Guarantees  
<http://proceedings.mlr.press/v97/gimenez19a.html>  
AUTHORS: Jaime Roquero Gimenez, James Zou  
HIGHLIGHT: We study a more fine-grained statistical problem: conditional feature selection, where a feature may be relevant depending on the values of the other features.

232, TITLE: Estimating Information Flow in Deep Neural Networks  
<http://proceedings.mlr.press/v97/goldfeld19a.html>  
AUTHORS: Ziv Goldfeld, Ewout Van Den Berg, Kristjan Greenewald, Igor Melnyk, Nam Nguyen, Brian Kingsbury, Yury Polyanskiy  
HIGHLIGHT: Focusing on feedforward networks with fixed weights and noisy internal representations, we develop a rigorous framework for accurate estimation of  $I(X; T_{\ell})$ .

233, TITLE: Amortized Monte Carlo Integration  
<http://proceedings.mlr.press/v97/golinski19a.html>  
AUTHORS: Adam Golinski, Frank Wood, Tom Rainforth  
HIGHLIGHT: In this paper, we address this inefficiency by introducing AMCI, a method for amortizing Monte Carlo integration directly.

234, TITLE: Online Algorithms for Rent-Or-Buy with Expert Advice  
<http://proceedings.mlr.press/v97/gollapudi19a.html>  
AUTHORS: Sreenivas Gollapudi, Debmalya Panigrahi  
HIGHLIGHT: In particular, we consider the classical rent-or-buy problem (also called ski rental), and obtain algorithms that provably improve their performance over the adversarial scenario by using these predictions.

235, TITLE: The information-theoretic value of unlabeled data in semi-supervised learning  
<http://proceedings.mlr.press/v97/golovnev19a.html>  
AUTHORS: Alexander Golovnev, David Pal, Balazs Szorenyi  
HIGHLIGHT: More specifically, we prove a separation by  $\Theta(\log n)$  multiplicative factor for the class of projections over the Boolean hypercube of dimension  $n$ .

236, TITLE: Efficient Training of BERT by Progressively Stacking  
<http://proceedings.mlr.press/v97/gong19a.html>  
AUTHORS: Linyuan Gong, Di He, Zhuohan Li, Tao Qin, Liwei Wang, Tiejian Liu  
HIGHLIGHT: In this paper, we explore an efficient training method for the state-of-the-art bidirectional Transformer (BERT) model.

- 237, TITLE: Quantile Stein Variational Gradient Descent for Batch Bayesian Optimization  
<http://proceedings.mlr.press/v97/gong19b.html>  
AUTHORS: Chengyue Gong, Jian Peng, Qiang Liu  
HIGHLIGHT: In this paper, we introduce a novel variational framework for batch query optimization, based on the argument that the query batch should be selected to have both high diversity and good worst case performance.
- 238, TITLE: Obtaining Fairness using Optimal Transport Theory  
<http://proceedings.mlr.press/v97/gordaliza19a.html>  
AUTHORS: Paula Gordaliza, Eustasio Del Barrio, Gamboa Fabrice, Loubes Jean-Michel  
HIGHLIGHT: We propose a Random Repair which yields a tradeoff between minimal information loss and a certain amount of fairness.
- 239, TITLE: Combining parametric and nonparametric models for off-policy evaluation  
<http://proceedings.mlr.press/v97/gottesman19a.html>  
AUTHORS: Omer Gottesman, Yao Liu, Scott Sussex, Emma Brunskill, Finale Doshi-Velez  
HIGHLIGHT: We consider a model-based approach to perform batch off-policy evaluation in reinforcement learning.
- 240, TITLE: Counterfactual Visual Explanations  
<http://proceedings.mlr.press/v97/goyal19a.html>  
AUTHORS: Yash Goyal, Ziyang Wu, Jan Ernst, Dhruv Batra, Devi Parikh, Stefan Lee  
HIGHLIGHT: In this work, we develop a technique to produce counterfactual visual explanations.
- 241, TITLE: Adaptive Sensor Placement for Continuous Spaces  
<http://proceedings.mlr.press/v97/grant19a.html>  
AUTHORS: James Grant, Alexis Boukouvalas, Ryan-Rhys Griffiths, David Leslie, Sattar Vakili, Enrique Munoz De Cote  
HIGHLIGHT: We present a new formulation of the problem as a continuum-armed bandit problem with feedback in the form of partial observations of realisations of an inhomogeneous Poisson process.
- 242, TITLE: A Statistical Investigation of Long Memory in Language and Music  
<http://proceedings.mlr.press/v97/greaves-tunnell19a.html>  
AUTHORS: Alexander Greaves-Tunnell, Zaid Harchaoui  
HIGHLIGHT: We contribute a statistical framework for investigating long-range dependence in current applications of deep sequence modeling, drawing on the well-developed theory of long memory stochastic processes.
- 243, TITLE: Automatic Posterior Transformation for Likelihood-Free Inference  
<http://proceedings.mlr.press/v97/greenberg19a.html>  
AUTHORS: David Greenberg, Marcel Nonnenmacher, Jakob Macke  
HIGHLIGHT: Here we present automatic posterior transformation (APT), a new sequential neural posterior estimation method for simulation-based inference.
- 244, TITLE: Learning to Optimize Multigrid PDE Solvers  
<http://proceedings.mlr.press/v97/greenfeld19a.html>  
AUTHORS: Daniel Greenfeld, Meirav Galun, Ronen Basri, Irad Yavneh, Ron Kimmel  
HIGHLIGHT: In this paper we propose a framework for learning multigrid solvers.
- 245, TITLE: Multi-Object Representation Learning with Iterative Variational Inference  
<http://proceedings.mlr.press/v97/greff19a.html>  
AUTHORS: Klaus Greff, Raphaël Lopez Kaufman, Rishabh Kabra, Nick Watters, Christopher Burgess, Daniel Zoran, Loic Matthey, Matthew Botvinick, Alexander Lerchner  
HIGHLIGHT: Instead, we argue for the importance of learning to segment and represent objects jointly.

- 246, TITLE: Graphite: Iterative Generative Modeling of Graphs  
<http://proceedings.mlr.press/v97/grover19a.html>  
AUTHORS: Aditya Grover, Aaron Zweig, Stefano Ermon  
HIGHLIGHT: In this work, we propose Graphite, an algorithmic framework for unsupervised learning of representations over nodes in large graphs using deep latent variable generative models.
- 247, TITLE: Fast Algorithm for Generalized Multinomial Models with Ranking Data  
<http://proceedings.mlr.press/v97/gu19a.html>  
AUTHORS: Jiaqi Gu, Guosheng Yin  
HIGHLIGHT: Based on this property, we propose an iterative algorithm that is easy to implement and interpret, and is guaranteed to converge.
- 248, TITLE: Towards a Deep and Unified Understanding of Deep Neural Models in NLP  
<http://proceedings.mlr.press/v97/guan19a.html>  
AUTHORS: Chaoyu Guan, Xiting Wang, Quanshi Zhang, Runjin Chen, Di He, Xing Xie  
HIGHLIGHT: We define a unified information-based measure to provide quantitative explanations on how intermediate layers of deep Natural Language Processing (NLP) models leverage information of input words.
- 249, TITLE: An Investigation of Model-Free Planning  
<http://proceedings.mlr.press/v97/guez19a.html>  
AUTHORS: Arthur Guez, Mehdi Mirza, Karol Gregor, Rishabh Kabra, Sebastien Racaniere, Theophane Weber, David Raposo, Adam Santoro, Laurent Orseau, Tom Eccles, Greg Wayne, David Silver, Timothy Lillicrap  
HIGHLIGHT: In this paper, we go even further, and demonstrate empirically that an entirely model-free approach, without special structure beyond standard neural network components such as convolutional networks and LSTMs, can learn to exhibit many of the characteristics typically associated with a model-based planner.
- 250, TITLE: Humor in Word Embeddings: Cockamamie Gobbledegook for Nincompoops  
<http://proceedings.mlr.press/v97/gultchin19a.html>  
AUTHORS: Limor Gultchin, Genevieve Patterson, Nancy Baym, Nathaniel Swinger, Adam Kalai  
HIGHLIGHT: While humor is often thought to be beyond the reach of Natural Language Processing, we show that several aspects of single-word humor correlate with simple linear directions in Word Embeddings.
- 251, TITLE: Simple Black-box Adversarial Attacks  
<http://proceedings.mlr.press/v97/guo19a.html>  
AUTHORS: Chuan Guo, Jacob Gardner, Yurong You, Andrew Gordon Wilson, Kilian Weinberger  
HIGHLIGHT: We propose an intriguingly simple method for the construction of adversarial images in the black-box setting.
- 252, TITLE: Exploring interpretable LSTM neural networks over multi-variable data  
<http://proceedings.mlr.press/v97/guo19b.html>  
AUTHORS: Tian Guo, Tao Lin, Nino Antulov-Fantulin  
HIGHLIGHT: In this paper, we explore the structure of LSTM recurrent neural networks to learn variable-wise hidden states, with the aim to capture different dynamics in multi-variable time series and distinguish the contribution of variables to the prediction.
- 253, TITLE: Learning to Exploit Long-term Relational Dependencies in Knowledge Graphs  
<http://proceedings.mlr.press/v97/guo19c.html>  
AUTHORS: Lingbing Guo, Zequn Sun, Wei Hu  
HIGHLIGHT: In this paper, we propose recurrent skipping networks (RSNs), which employ a skipping mechanism to bridge the gaps between entities.

- 254, TITLE: Memory-Optimal Direct Convolutions for Maximizing Classification Accuracy in Embedded Applications  
http://proceedings.mlr.press/v97/gural19a.html  
AUTHORS: Albert Gural, Boris Murmann  
HIGHLIGHT: This paper presents memory-optimal direct convolutions as a way to push classification accuracy as high as possible given strict hardware memory constraints at the expense of extra compute.
- 255, TITLE: IMEXnet A Forward Stable Deep Neural Network  
http://proceedings.mlr.press/v97/haber19a.html  
AUTHORS: Eldad Haber, Keegan Lensink, Eran Treister, Lars Ruthotto  
HIGHLIGHT: We introduce the IMEXnet that addresses these challenges by adapting semi-implicit methods for partial differential equations.  
We also present a new dataset for semantic segmentation and demonstrate the effectiveness of our architecture using the NYU Depth dataset.
- 256, TITLE: On The Power of Curriculum Learning in Training Deep Networks  
http://proceedings.mlr.press/v97/hacohen19a.html  
AUTHORS: Guy Hacohen, Daphna Weinshall  
HIGHLIGHT: In this work, we analyze the effect of curriculum learning, which involves the non-uniform sampling of mini-batches, on the training of deep networks, and specifically CNNs trained for image recognition.
- 257, TITLE: Trading Redundancy for Communication: Speeding up Distributed SGD for Non-convex Optimization  
http://proceedings.mlr.press/v97/haddadpour19a.html  
AUTHORS: Farzin Haddadpour, Mohammad Mahdi Kamani, Mehrdad Mahdavi, Viveck Cadambe  
HIGHLIGHT: In this paper, we advocate the use of redundancy towards communication-efficient distributed stochastic algorithms for non-convex optimization.
- 258, TITLE: Learning Latent Dynamics for Planning from Pixels  
http://proceedings.mlr.press/v97/hafner19a.html  
AUTHORS: Danijar Hafner, Timothy Lillicrap, Ian Fischer, Ruben Villegas, David Ha, Honglak Lee, James Davidson  
HIGHLIGHT: We propose the Deep Planning Network (PlaNet), a purely model-based agent that learns the environment dynamics from images and chooses actions through fast online planning in latent space.
- 259, TITLE: Neural Separation of Observed and Unobserved Distributions  
http://proceedings.mlr.press/v97/halperin19a.html  
AUTHORS: Tavi Halperin, Ariel Ephrat, Yedid Hoshen  
HIGHLIGHT: In this work, we introduce a new method—Neural Egg Separation—to tackle the scenario of extracting a signal from an unobserved distribution additively mixed with a signal from an observed distribution.
- 260, TITLE: Grid-Wise Control for Multi-Agent Reinforcement Learning in Video Game AI  
http://proceedings.mlr.press/v97/han19a.html  
AUTHORS: Lei Han, Peng Sun, Yali Du, Jiechao Xiong, Qing Wang, Xinghai Sun, Han Liu, Tong Zhang  
HIGHLIGHT: To address the issue, we propose a novel architecture that learns a spatial joint representation of all the agents and outputs grid-wise actions.
- 261, TITLE: Dimension-Wise Importance Sampling Weight Clipping for Sample-Efficient Reinforcement Learning  
http://proceedings.mlr.press/v97/han19b.html  
AUTHORS: Seungyul Han, Youngchul Sung  
HIGHLIGHT: In this paper, we consider PPO, a representative on-policy algorithm, and propose its improvement by dimension-wise IS weight clipping which separately clips the IS weight of each action dimension to avoid large bias and adaptively controls the IS weight to bound policy update from the current policy.

- 262, TITLE: Complexity of Linear Regions in Deep Networks  
<http://proceedings.mlr.press/v97/hanin19a.html>  
AUTHORS: Boris Hanin, David Rolnick  
HIGHLIGHT: In this paper, we provide a mathematical framework to count the number of linear regions of a piecewise linear network and measure the volume of the boundaries between these regions.
- 263, TITLE: Importance Sampling Policy Evaluation with an Estimated Behavior Policy  
<http://proceedings.mlr.press/v97/hanna19a.html>  
AUTHORS: Josiah Hanna, Scott Niekum, Peter Stone  
HIGHLIGHT: In this paper, we study importance sampling with an estimated behavior policy where the behavior policy estimate comes from the same set of data used to compute the importance sampling estimate.
- 264, TITLE: Doubly-Competitive Distribution Estimation  
<http://proceedings.mlr.press/v97/hao19a.html>  
AUTHORS: Yi Hao, Alon Orlitsky  
HIGHLIGHT: This paper combines and strengthens the two frameworks.
- 265, TITLE: Random Shuffling Beats SGD after Finite Epochs  
<http://proceedings.mlr.press/v97/haochen19a.html>  
AUTHORS: Jeff Haochen, Suvrit Sra  
HIGHLIGHT: Building upon [gurbuzbalaban2015random](#), we present the first (to our knowledge) non-asymptotic results for this problem by proving that after a reasonable number of epochs  $\text{rsgd}$  converges faster than  $\text{sgd}$ .
- 266, TITLE: Submodular Maximization beyond Non-negativity: Guarantees, Fast Algorithms, and Applications  
<http://proceedings.mlr.press/v97/harshaw19a.html>  
AUTHORS: Chris Harshaw, Moran Feldman, Justin Ward, Amin Karbasi  
HIGHLIGHT: We present an algorithm for maximizing  $g - c$  under a  $k$ -cardinality constraint which produces a random feasible set  $S$  such that  $\mathbb{E}[g(S) - c(S)] \geq (1 - e^{-\gamma} - \epsilon) g(\text{opt}) - c(\text{opt})$ , whose running time is  $\mathcal{O}(\frac{1}{\epsilon} \log^2 \frac{1}{\epsilon})$ , independent of  $k$ .
- 267, TITLE: Per-Decision Option Discounting  
<http://proceedings.mlr.press/v97/harutyunyan19a.html>  
AUTHORS: Anna Harutyunyan, Peter Vrancx, Philippe Hamel, Ann Nowe, Doina Precup  
HIGHLIGHT: We propose a modification to the options framework that naturally scales the agent's horizon with option length.
- 268, TITLE: Submodular Observation Selection and Information Gathering for Quadratic Models  
<http://proceedings.mlr.press/v97/hashemi19a.html>  
AUTHORS: Abolfazl Hashemi, Mahsa Ghasemi, Haris Vikalo, Ufuk Topcu  
HIGHLIGHT: We study the problem of selecting most informative subset of a large observation set to enable accurate estimation of unknown parameters.
- 269, TITLE: Understanding and Controlling Memory in Recurrent Neural Networks  
<http://proceedings.mlr.press/v97/haviv19a.html>  
AUTHORS: Doron Haviv, Alexander Rivkind, Omri Barak  
HIGHLIGHT: Here, we utilize different training protocols, datasets and architectures to obtain a range of networks solving a delayed classification task with similar performance, alongside substantial differences in their ability to extrapolate for longer delays.
- 270, TITLE: On the Impact of the Activation function on Deep Neural Networks Training  
<http://proceedings.mlr.press/v97/hayou19a.html>  
AUTHORS: Soufiane Hayou, Arnaud Doucet, Judith Rousseau



**HIGHLIGHT:** We give a comprehensive theoretical analysis of the Edge of Chaos and show that we can indeed tune the initialization parameters and the activation function in order to accelerate the training and improve the performance.

271, **TITLE:** Provably Efficient Maximum Entropy Exploration

<http://proceedings.mlr.press/v97/hazan19a.html>

**AUTHORS:** Elad Hazan, Sham Kakade, Karan Singh, Abby Van Soest

**HIGHLIGHT:** We provide an efficient algorithm to optimize such such intrinsically defined objectives, when given access to a black box planning oracle (which is robust to function approximation).

272, **TITLE:** On the Long-term Impact of Algorithmic Decision Policies: Effort Unfairness and Feature Segregation through Social Learning

<http://proceedings.mlr.press/v97/heidari19a.html>

**AUTHORS:** Hoda Heidari, Vedant Nanda, Krishna Gummadi

**HIGHLIGHT:** We propose an effort-based measure of fairness and present a data-driven framework for characterizing the long-term impact of algorithmic policies on reshaping the underlying population.

273, **TITLE:** Graph Resistance and Learning from Pairwise Comparisons

<http://proceedings.mlr.press/v97/hendrickx19a.html>

**AUTHORS:** Julien Hendrickx, Alexander Olshevsky, Venkatesh Saligrama

**HIGHLIGHT:** We consider the problem of learning the qualities of a collection of items by performing noisy comparisons among them.

274, **TITLE:** Using Pre-Training Can Improve Model Robustness and Uncertainty

<http://proceedings.mlr.press/v97/hendrycks19a.html>

**AUTHORS:** Dan Hendrycks, Kimin Lee, Mantas Mazeika

**HIGHLIGHT:** We show that although pre-training may not improve performance on traditional classification metrics, it improves model robustness and uncertainty estimates.

275, **TITLE:** Flow++: Improving Flow-Based Generative Models with Variational Dequantization and Architecture Design

<http://proceedings.mlr.press/v97/ho19a.html>

**AUTHORS:** Jonathan Ho, Xi Chen, Aravind Srinivas, Yan Duan, Pieter Abbeel

**HIGHLIGHT:** In this paper, we investigate and improve upon three limiting design choices employed by flow-based models in prior work: the use of uniform noise for dequantization, the use of inexpressive affine flows, and the use of purely convolutional conditioning networks in coupling layers.

276, **TITLE:** Population Based Augmentation: Efficient Learning of Augmentation Policy Schedules

<http://proceedings.mlr.press/v97/ho19b.html>

**AUTHORS:** Daniel Ho, Eric Liang, Xi Chen, Ion Stoica, Pieter Abbeel

**HIGHLIGHT:** In this paper, we introduce a new data augmentation algorithm, Population Based Augmentation (PBA), which generates nonstationary augmentation policy schedules instead of a fixed augmentation policy.

277, **TITLE:** Collective Model Fusion for Multiple Black-Box Experts

<http://proceedings.mlr.press/v97/hoang19a.html>

**AUTHORS:** Minh Hoang, Nghia Hoang, Bryan Kian Hsiang Low, Carleton Kingsford

**HIGHLIGHT:** The proposed method will enable this by addressing the key issues of how black-box experts interact to understand the predictive behaviors of one another; how these understandings can be represented and shared efficiently among themselves; and how the shared understandings can be combined to generate high-quality consensus prediction.

278, **TITLE:** Connectivity-Optimized Representation Learning via Persistent Homology

<http://proceedings.mlr.press/v97/hofer19a.html>

**AUTHORS:** Christoph Hofer, Roland Kwitt, Marc Niethammer, Mandar Dixit

**HIGHLIGHT:** Under mild conditions, this loss is differentiable and we present a theoretical analysis of the properties induced by the loss.

279, **TITLE:** Better generalization with less data using robust gradient descent

<http://proceedings.mlr.press/v97/holland19a.html>

**AUTHORS:** Matthew Holland, Kazushi Ikeda

**HIGHLIGHT:** In pursuit of stronger performance under weaker assumptions, we propose a technique which uses a cheap and robust iterative estimate of the risk gradient, which can be easily fed into any steepest descent procedure.

280, **TITLE:** Emerging Convolutions for Generative Normalizing Flows

<http://proceedings.mlr.press/v97/hoogetboom19a.html>

**AUTHORS:** Emiel Hoogetboom, Rianne Van Den Berg, Max Welling

**HIGHLIGHT:** We propose two methods to produce invertible convolutions, that have receptive fields identical to standard convolutions: Emerging convolutions are obtained by chaining specific autoregressive convolutions, and periodic convolutions are decoupled in the frequency domain.

281, **TITLE:** Nonconvex Variance Reduced Optimization with Arbitrary Sampling

<http://proceedings.mlr.press/v97/horvath19a.html>

**AUTHORS:** Samuel Horvath, Peter Richtarik

**HIGHLIGHT:** We provide the first importance sampling variants of variance reduced algorithms for empirical risk minimization with non-convex loss functions.

282, **TITLE:** Parameter-Efficient Transfer Learning for NLP

<http://proceedings.mlr.press/v97/houlsby19a.html>

**AUTHORS:** Neil Houlsby, Andrei Giurgiu, Stanislaw Jastrzebski, Bruna Morrone, Quentin De Laroussilhe, Andrea Gesmundo, Mona Attariyan, Sylvain Gelly

**HIGHLIGHT:** As an alternative, we propose transfer with adapter modules.

283, **TITLE:** Stay With Me: Lifetime Maximization Through Heteroscedastic Linear Bandits With Reneging

<http://proceedings.mlr.press/v97/hsieh19a.html>

**AUTHORS:** Ping-Chun Hsieh, Xi Liu, Anirban Bhattacharya, P R Kumar

**HIGHLIGHT:** To address the above issue, this paper proposes a model of heteroscedastic linear bandits with reneing, which allows each participant to have a distinct "satisfaction level," with any interaction outcome falling short of that level resulting in that participant reneing.

284, **TITLE:** Finding Mixed Nash Equilibria of Generative Adversarial Networks

<http://proceedings.mlr.press/v97/hsieh19b.html>

**AUTHORS:** Ya-Ping Hsieh, Chen Liu, Volkan Cevher

**HIGHLIGHT:** In this work, we tackle the training of GANs by rethinking the problem formulation from the mixed Nash Equilibria (NE) perspective.

285, **TITLE:** Classification from Positive, Unlabeled and Biased Negative Data

<http://proceedings.mlr.press/v97/hsieh19c.html>

**AUTHORS:** Yu-Guan Hsieh, Gang Niu, Masashi Sugiyama

**HIGHLIGHT:** We provide a method based on empirical risk minimization to address this PUBN classification problem.

286, **TITLE:** Bayesian Deconditional Kernel Mean Embeddings

<http://proceedings.mlr.press/v97/hsu19a.html>

**AUTHORS:** Kelvin Hsu, Fabio Ramos

**HIGHLIGHT:** Critically, we introduce the notion of task transformed Gaussian processes and establish deconditional kernel means embeddings as their posterior predictive mean.

- 287, TITLE: Faster Stochastic Alternating Direction Method of Multipliers for Nonconvex Optimization  
<http://proceedings.mlr.press/v97/huang19a.html>  
AUTHORS: Feihu Huang, Songcan Chen, Heng Huang  
HIGHLIGHT: In this paper, we propose a faster stochastic alternating direction method of multipliers (ADMM) for nonconvex optimization by using a new stochastic path-integrated differential estimator (SPIDER), called as SPIDER-ADMM.
- 288, TITLE: Unsupervised Deep Learning by Neighbourhood Discovery  
<http://proceedings.mlr.press/v97/huang19b.html>  
AUTHORS: Jiabo Huang, Qi Dong, Shaogang Gong, Xiatian Zhu  
HIGHLIGHT: In this work, we introduce a generic unsupervised deep learning approach to training deep models without the need for any manual label supervision.
- 289, TITLE: Detecting Overlapping and Correlated Communities without Pure Nodes: Identifiability and Algorithm  
<http://proceedings.mlr.press/v97/huang19c.html>  
AUTHORS: Kejun Huang, Xiao Fu  
HIGHLIGHT: We adopt the mixed-membership stochastic blockmodel as the underlying probabilistic model, and give conditions under which the memberships of a subset of nodes can be uniquely identified.
- 290, TITLE: Hierarchical Importance Weighted Autoencoders  
<http://proceedings.mlr.press/v97/huang19d.html>  
AUTHORS: Chin-Wei Huang, Kris Sankaran, Eeshan Dhekane, Alexandre Lacoste, Aaron Courville  
HIGHLIGHT: Theoretically, we analyze the condition under which convergence of the estimator variance can be connected to convergence of the lower bound.
- 291, TITLE: Stable and Fair Classification  
<http://proceedings.mlr.press/v97/huang19e.html>  
AUTHORS: Lingxiao Huang, Nisheeth Vishnoi  
HIGHLIGHT: We propose an extended framework based on fair classification algorithms that are formulated as optimization problems, by introducing a stability-focused regularization term.
- 292, TITLE: Addressing the Loss-Metric Mismatch with Adaptive Loss Alignment  
<http://proceedings.mlr.press/v97/huang19f.html>  
AUTHORS: Chen Huang, Shuangfei Zhai, Walter Talbott, Miguel Bautista Martin, Shih-Yu Sun, Carlos Guestrin, Josh Susskind  
HIGHLIGHT: In this work we assess this assumption by meta-learning an adaptive loss function to directly optimize the evaluation metric.
- 293, TITLE: Causal Discovery and Forecasting in Nonstationary Environments with State-Space Models  
<http://proceedings.mlr.press/v97/huang19g.html>  
AUTHORS: Biwei Huang, Kun Zhang, Mingming Gong, Clark Glymour  
HIGHLIGHT: In this paper, we study causal discovery and forecasting for nonstationary time series.
- 294, TITLE: Composing Entropic Policies using Divergence Correction  
<http://proceedings.mlr.press/v97/hunt19a.html>  
AUTHORS: Jonathan Hunt, Andre Barreto, Timothy Lillicrap, Nicolas Heess  
HIGHLIGHT: As part of this analysis, we extend an important generalization of policy improvement to the maximum entropy framework and introduce an algorithm for the practical implementation of successor features in continuous action spaces.
- 295, TITLE: HexaGAN: Generative Adversarial Nets for Real World Classification  
<http://proceedings.mlr.press/v97/hwang19a.html>  
AUTHORS: Uiwon Hwang, Dahuin Jung, Sungroh Yoon

**HIGHLIGHT:** In this paper, we propose HexaGAN, a generative adversarial network framework that shows promising classification performance for all three problems.

296, **TITLE:** Overcoming Mean-Field Approximations in Recurrent Gaussian Process Models  
<http://proceedings.mlr.press/v97/ialongo19a.html>  
**AUTHORS:** Alessandro Davide Ialongo, Mark Van Der Wilk, James Hensman, Carl Edward Rasmussen  
**HIGHLIGHT:** We identify a new variational inference scheme for dynamical systems whose transition function is modelled by a Gaussian process.

297, **TITLE:** Learning Structured Decision Problems with Unawareness  
<http://proceedings.mlr.press/v97/innes19a.html>  
**AUTHORS:** Craig Innes, Alex Lascarides  
**HIGHLIGHT:** In this paper, we learn Bayesian Decision Networks from both domain exploration and expert assertions in a way which guarantees convergence to optimal behaviour, even when the agent starts unaware of actions or belief variables that are critical to success.

298, **TITLE:** Phase transition in PCA with missing data: Reduced signal-to-noise ratio, not sample size!  
<http://proceedings.mlr.press/v97/ipsen19a.html>  
**AUTHORS:** Niels Ipsen, Lars Kai Hansen  
**HIGHLIGHT:** Here we generalize this analysis to include missing data.

299, **TITLE:** Actor-Attention-Critic for Multi-Agent Reinforcement Learning  
<http://proceedings.mlr.press/v97/iqbal19a.html>  
**AUTHORS:** Shariq Iqbal, Fei Sha  
**HIGHLIGHT:** We present an actor-critic algorithm that trains decentralized policies in multi-agent settings, using centrally computed critics that share an attention mechanism which selects relevant information for each agent at every timestep.

300, **TITLE:** Complementary-Label Learning for Arbitrary Losses and Models  
<http://proceedings.mlr.press/v97/ishida19a.html>  
**AUTHORS:** Takashi Ishida, Gang Niu, Aditya Menon, Masashi Sugiyama  
**HIGHLIGHT:** The goal of this paper is to derive a novel framework of complementary-label learning with an unbiased estimator of the classification risk, for arbitrary losses and models—all existing methods have failed to achieve this goal.

301, **TITLE:** Causal Identification under Markov Equivalence: Completeness Results  
<http://proceedings.mlr.press/v97/jaber19a.html>  
**AUTHORS:** Amin Jaber, Jiji Zhang, Elias Bareinboim  
**HIGHLIGHT:** In this paper, we relax this requirement and consider that the knowledge is articulated in the form of an equivalence class of causal diagrams, in particular, a partial ancestral graph (PAG).

302, **TITLE:** Learning from a Learner  
<http://proceedings.mlr.press/v97/jacq19a.html>  
**AUTHORS:** Alexis Jacq, Matthieu Geist, Ana Paiva, Olivier Pietquin  
**HIGHLIGHT:** In this paper, we propose a novel setting for Inverse Reinforcement Learning (IRL), namely "Learning from a Learner" (L<sub>L</sub>).

303, **TITLE:** Differentially Private Fair Learning  
<http://proceedings.mlr.press/v97/jagielski19a.html>  
**AUTHORS:** Matthew Jagielski, Michael Kearns, Jieming Mao, Alina Oprea, Aaron Roth, Saeed Sharifi - Malvajerdi, Jonathan Ullman  
**HIGHLIGHT:** Motivated by settings in which predictive models may be required to be non-discriminatory with respect to certain attributes (such as race), but even collecting the sensitive attribute may be forbidden or restricted, we initiate the study of fair learning under the constraint of differential privacy.

- 304, TITLE: Sum-of-Squares Polynomial Flow  
<http://proceedings.mlr.press/v97/jaini19a.html>  
AUTHORS: Priyank Jaini, Kira A. Selby, Yaoliang Yu  
HIGHLIGHT: Based on triangular maps, we propose a general framework for high-dimensional density estimation, by specifying one-dimensional transformations (equivalently conditional densities) and appropriate conditioner networks.
- 305, TITLE: DBSCAN++: Towards fast and scalable density clustering  
<http://proceedings.mlr.press/v97/jang19a.html>  
AUTHORS: Jennifer Jang, Heinrich Jiang  
HIGHLIGHT: We propose DBSCAN++, a simple modification of DBSCAN which only requires computing the densities for a chosen subset of points.
- 306, TITLE: Learning What and Where to Transfer  
<http://proceedings.mlr.press/v97/jang19b.html>  
AUTHORS: Yunhun Jang, Hankook Lee, Sung Ju Hwang, Jinwoo Shin  
HIGHLIGHT: To address the issue, we propose a novel transfer learning approach based on meta-learning that can automatically learn what knowledge to transfer from the source network to where in the target network.
- 307, TITLE: Social Influence as Intrinsic Motivation for Multi-Agent Deep Reinforcement Learning  
<http://proceedings.mlr.press/v97/jaques19a.html>  
AUTHORS: Natasha Jaques, Angeliki Lazaridou, Edward Hughes, Caglar Gulcehre, Pedro Ortega, Dj Strouse, Joel Z Leibo, Nando De Freitas  
HIGHLIGHT: We propose a unified mechanism for achieving coordination and communication in Multi-Agent Reinforcement Learning (MARL), through rewarding agents for having causal influence over other agents' actions.
- 308, TITLE: A Deep Reinforcement Learning Perspective on Internet Congestion Control  
<http://proceedings.mlr.press/v97/jay19a.html>  
AUTHORS: Nathan Jay, Noga Rotman, Brighten Godfrey, Michael Schapira, Aviv Tamar  
HIGHLIGHT: We present and investigate a novel and timely application domain for deep reinforcement learning (RL): Internet congestion control.
- 309, TITLE: Graph Neural Network for Music Score Data and Modeling Expressive Piano Performance  
<http://proceedings.mlr.press/v97/jeong19a.html>  
AUTHORS: Dasaem Jeong, Taegyun Kwon, Yoojin Kim, Juhan Nam  
HIGHLIGHT: In this paper, we represent the unique form of musical score using graph neural network and apply it for rendering expressive piano performance from the music score.
- 310, TITLE: Ladder Capsule Network  
<http://proceedings.mlr.press/v97/jeong19b.html>  
AUTHORS: Taewon Jeong, Youngmin Lee, Heeyoung Kim  
HIGHLIGHT: We propose a new architecture of the capsule network called the ladder capsule network, which has an alternative building block to the dynamic routing algorithm in the capsule network (Sabour et al., 2017).
- 311, TITLE: Training CNNs with Selective Allocation of Channels  
<http://proceedings.mlr.press/v97/jeong19c.html>  
AUTHORS: Jongheon Jeong, Jinwoo Shin  
HIGHLIGHT: In this paper, we propose a simple way to improve the capacity of any CNN model having large-scale features, without adding more parameters.
- 312, TITLE: Learning Discrete and Continuous Factors of Data via Alternating Disentanglement  
<http://proceedings.mlr.press/v97/jeong19d.html>

- AUTHORS: Yeonwoo Jeong, Hyun Oh Song  
HIGHLIGHT: We address the problem of unsupervised disentanglement of discrete and continuous explanatory factors of data.
- 313, TITLE: Improved Zeroth-Order Variance Reduced Algorithms and Analysis for Nonconvex Optimization  
<http://proceedings.mlr.press/v97/ji19a.html>  
AUTHORS: Kaiyi Ji, Zhe Wang, Yi Zhou, Yingbin Liang  
HIGHLIGHT: In this paper, we propose a new algorithm ZO-SVRG-Coord-Rand and develop a new analysis for an existing ZO-SVRG-Coord algorithm proposed in Liu et al. 2018b, and show that both ZO-SVRG-Coord-Rand and ZO-SVRG-Coord (under our new analysis) outperform other exiting SVRG-type zeroth-order methods as well as ZO-GD and ZO-SGD.
- 314, TITLE: Neural Logic Reinforcement Learning  
<http://proceedings.mlr.press/v97/jiang19a.html>  
AUTHORS: Zhengyao Jiang, Shan Luo  
HIGHLIGHT: To address these two challenges, we propose a novel algorithm named Neural Logic Reinforcement Learning (NLRL) to represent the policies in reinforcement learning by first-order logic.
- 315, TITLE: Finding Options that Minimize Planning Time  
<http://proceedings.mlr.press/v97/jinnai19a.html>  
AUTHORS: Yuu Jinnai, David Abel, David Hershkowitz, Michael Littman, George Konidaris  
HIGHLIGHT: We formalize the problem of selecting the optimal set of options for planning as that of computing the smallest set of options so that planning converges in less than a given maximum of value-iteration passes.
- 316, TITLE: Discovering Options for Exploration by Minimizing Cover Time  
<http://proceedings.mlr.press/v97/jinnai19b.html>  
AUTHORS: Yuu Jinnai, Jee Won Park, David Abel, George Konidaris  
HIGHLIGHT: We introduce a new option discovery algorithm that diminishes the expected cover time by connecting the most distant states in the state-space graph with options.
- 317, TITLE: Kernel Mean Matching for Content Addressability of GANs  
<http://proceedings.mlr.press/v97/jitkrittum19a.html>  
AUTHORS: Wittawat Jitkrittum, Patsorn Sangkloy, Muhammad Waleed Gondal, Amit Raj, James Hays, Bernhard Schölkopf  
HIGHLIGHT: We propose a novel procedure which adds "content-addressability" to any given unconditional implicit model e.g., a generative adversarial network (GAN).
- 318, TITLE: GOODE: A Gaussian Off-The-Shelf Ordinary Differential Equation Solver  
<http://proceedings.mlr.press/v97/john19a.html>  
AUTHORS: David John, Vincent Heuveline, Michael Schober  
HIGHLIGHT: Our method based on iterated Gaussian process (GP) regression returns a GP posterior over the solution of nonlinear ODEs, which provides a meaningful error estimation via its predictive posterior standard deviation.
- 319, TITLE: Bilinear Bandits with Low-rank Structure  
<http://proceedings.mlr.press/v97/jun19a.html>  
AUTHORS: Kwang-Sung Jun, Rebecca Willett, Stephen Wright, Robert Nowak  
HIGHLIGHT: We introduce the bilinear bandit problem with low-rank structure in which an action takes the form of a pair of arms from two different entity types, and the reward is a bilinear function of the known feature vectors of the arms.
- 320, TITLE: Statistical Foundations of Virtual Democracy  
<http://proceedings.mlr.press/v97/kahng19a.html>  
AUTHORS: Anson Kahng, Min Kyung Lee, Ritesh Noothigattu, Ariel Procaccia, Christos-Alexandros Psomas

**HIGHLIGHT:** One of the key questions is which aggregation method – or voting rule – to use; we offer a novel statistical viewpoint that provides guidance.

321, **TITLE:** Molecular Hypergraph Grammar with Its Application to Molecular Optimization

<http://proceedings.mlr.press/v97/kajino19a.html>

**AUTHORS:** Hiroshi Kajino

**HIGHLIGHT:** This paper presents a molecular hypergraph grammar variational autoencoder (MHG-VAE), which uses a single VAE to achieve 100% validity.

322, **TITLE:** Robust Influence Maximization for Hyperparametric Models

<http://proceedings.mlr.press/v97/kalimeris19a.html>

**AUTHORS:** Dimitris Kalimeris, Gal Kaplun, Yaron Singer

**HIGHLIGHT:** In this paper we study the problem of robust influence maximization in the independent cascade model under a hyperparametric assumption.

323, **TITLE:** Classifying Treatment Responders Under Causal Effect Monotonicity

<http://proceedings.mlr.press/v97/kallus19a.html>

**AUTHORS:** Nathan Kallus

**HIGHLIGHT:** In the context of individual-level causal inference, we study the problem of predicting whether someone will respond or not to a treatment based on their features and past examples of features, treatment indicator (e.g., drug/no drug), and a binary outcome (e.g., recovery from disease).

324, **TITLE:** Trainable Decoding of Sets of Sequences for Neural Sequence Models

<http://proceedings.mlr.press/v97/kalyan19a.html>

**AUTHORS:** Ashwin Kalyan, Peter Anderson, Stefan Lee, Dhruv Batra

**HIGHLIGHT:** To address this, we propose  $\nabla$ SBS, a trainable decoding procedure that outputs a set of sequences, highly valued according to the metric.

325, **TITLE:** Myopic Posterior Sampling for Adaptive Goal Oriented Design of Experiments

<http://proceedings.mlr.press/v97/kandasamy19a.html>

**AUTHORS:** Kirthivasan Kandasamy, Willie Neiswanger, Reed Zhang, Akshay Krishnamurthy, Jeff Schneider, Barnabas Poczos

**HIGHLIGHT:** In this work, we design a new myopic strategy for a wide class of adaptive design of experiment (DOE) problems, where we wish to collect data in order to fulfil a given goal.

326, **TITLE:** Differentially Private Learning of Geometric Concepts

<http://proceedings.mlr.press/v97/kaplan19a.html>

**AUTHORS:** Haim Kaplan, Yishay Mansour, Yossi Matias, Uri Stemmer

**HIGHLIGHT:** We present differentially private efficient algorithms for learning union of polygons in the plane (which are not necessarily convex).

327, **TITLE:** Policy Consolidation for Continual Reinforcement Learning

<http://proceedings.mlr.press/v97/kaplanis19a.html>

**AUTHORS:** Christos Kaplanis, Murray Shanahan, Claudia Clopath

**HIGHLIGHT:** We propose a method for tackling catastrophic forgetting in deep reinforcement learning that is agnostic to the timescale of changes in the distribution of experiences, does not require knowledge of task boundaries and can adapt in continuously changing environments.

328, **TITLE:** Error Feedback Fixes SignSGD and other Gradient Compression Schemes

<http://proceedings.mlr.press/v97/karimireddy19a.html>

**AUTHORS:** Sai Praneeth Karimireddy, Quentin Rebjock, Sebastian Stich, Martin Jaggi

**HIGHLIGHT:** We show simple convex counter-examples where signSGD does not converge to the optimum.

329, TITLE: Riemannian adaptive stochastic gradient algorithms on matrix manifolds  
<http://proceedings.mlr.press/v97/kasai19a.html>  
AUTHORS: Hiroyuki Kasai, Pratik Jawanpuria, Bamdev Mishra  
HIGHLIGHT: We propose novel stochastic gradient algorithms for problems on Riemannian matrix manifolds by adapting the row and column subspaces of gradients.

330, TITLE: Neural Inverse Knitting: From Images to Manufacturing Instructions  
<http://proceedings.mlr.press/v97/kaspar19a.html>  
AUTHORS: Alexandre Kaspar, Tae-Hyun Oh, Liane Makatura, Petr Kellnhofer, Wojciech Matusik  
HIGHLIGHT: Motivated by the recent potential of mass customization brought by whole-garment knitting machines, we introduce the new problem of automatic machine instruction generation using a single image of the desired physical product, which we apply to machine knitting.  
We create a curated dataset of real samples with their instruction counterpart and propose to use synthetic images to augment it in a novel way.

331, TITLE: Processing Megapixel Images with Deep Attention-Sampling Models  
<http://proceedings.mlr.press/v97/katharopoulos19a.html>  
AUTHORS: Angelos Katharopoulos, Francois Fleuret  
HIGHLIGHT: To tackle this limitation, we propose a fully differentiable end-to-end trainable model that samples and processes only a fraction of the full resolution input image.

332, TITLE: Robust Estimation of Tree Structured Gaussian Graphical Models  
<http://proceedings.mlr.press/v97/katiyar19a.html>  
AUTHORS: Ashish Katiyar, Jessica Hoffmann, Constantine Caramanis  
HIGHLIGHT: Robust Estimation of Tree Structured Gaussian Graphical Models.

333, TITLE: Shallow-Deep Networks: Understanding and Mitigating Network Overthinking  
<http://proceedings.mlr.press/v97/kaya19a.html>  
AUTHORS: Yigitcan Kaya, Sanghyun Hong, Tudor Dumitras  
HIGHLIGHT: For prediction transparency, we propose the Shallow-Deep Network (SDN), a generic modification to off-the-shelf DNNs that introduces internal classifiers.

334, TITLE: Submodular Streaming in All Its Glory: Tight Approximation, Minimum Memory and Low Adaptive Complexity  
<http://proceedings.mlr.press/v97/kazemi19a.html>  
AUTHORS: Ehsan Kazemi, Marko Mitrovic, Morteza Zadimoghaddam, Silvio Lattanzi, Amin Karbasi  
HIGHLIGHT: In this paper, we study the problem of maximizing a monotone submodular function in the streaming setting with a cardinality constraint  $k$ .

335, TITLE: Adaptive Scale-Invariant Online Algorithms for Learning Linear Models  
<http://proceedings.mlr.press/v97/kempka19a.html>  
AUTHORS: Michal Kempka, Wojciech Kotlowski, Manfred K. Warmuth  
HIGHLIGHT: In this paper, we resolve the tuning problem by proposing online algorithms making predictions which are invariant under arbitrary rescaling of the features.

336, TITLE: CHiVE: Varying Prosody in Speech Synthesis with a Linguistically Driven Dynamic Hierarchical Conditional Variational Network  
<http://proceedings.mlr.press/v97/kenter19a.html>  
AUTHORS: Tom Kenter, Vincent Wan, Chun-An Chan, Rob Clark, Jakub Vit  
HIGHLIGHT: We present a new, hierarchically structured conditional variational auto-encoder to generate prosodic features (fundamental frequency, energy and duration) suitable for use with a vocoder or a generative model like WaveNet.

337, TITLE: Collaborative Evolutionary Reinforcement Learning



<http://proceedings.mlr.press/v97/khadka19a.html>

AUTHORS: Shauharda Khadka, Somdeb Majumdar, Tarek Nassar, Zach Dwiell, Evren Tumer, Santiago Miret, Yinyin Liu, Kagan Tumer

HIGHLIGHT: In this paper, we introduce Collaborative Evolutionary Reinforcement Learning (CERL), a scalable framework that comprises a portfolio of policies that simultaneously explore and exploit diverse regions of the solution space.

338, TITLE: Geometry Aware Convolutional Filters for Omnidirectional Images Representation

<http://proceedings.mlr.press/v97/khasanova19a.html>

AUTHORS: Renata Khasanova, Pascal Frossard

HIGHLIGHT: In this paper we aim at improving popular deep convolutional neural networks so that they can properly take into account the specific properties of omnidirectional data.

339, TITLE: EMI: Exploration with Mutual Information

<http://proceedings.mlr.press/v97/kim19a.html>

AUTHORS: Hyoungseok Kim, Jaekyeom Kim, Yeonwoo Jeong, Sergey Levine, Hyun Oh Song

HIGHLIGHT: We propose EMI, which is an exploration method that constructs embedding representation of states and actions that does not rely on generative decoding of the full observation but extracts predictive signals that can be used to guide exploration based on forward prediction in the representation space.

340, TITLE: FloWaveNet : A Generative Flow for Raw Audio

<http://proceedings.mlr.press/v97/kim19b.html>

AUTHORS: Sungwon Kim, Sang-Gil Lee, Jongyoon Song, Jaehyeon Kim, Sungroh Yoon

HIGHLIGHT: We propose FloWaveNet, a flow-based generative model for raw audio synthesis.

341, TITLE: Curiosity-Bottleneck: Exploration By Distilling Task-Specific Novelty

<http://proceedings.mlr.press/v97/kim19c.html>

AUTHORS: Youngjin Kim, Wontae Nam, Hyunwoo Kim, Ji-Hoon Kim, Gunhee Kim

HIGHLIGHT: We introduce an information- theoretic exploration strategy named Curiosity-Bottleneck that distills task-relevant information from observation.

342, TITLE: Contextual Multi-armed Bandit Algorithm for Semiparametric Reward Model

<http://proceedings.mlr.press/v97/kim19d.html>

AUTHORS: Gi-Soo Kim, Myunghye Cho Paik

HIGHLIGHT: This paper proposes a new contextual MAB algorithm for a relaxed, semiparametric reward model that supports nonstationarity.

343, TITLE: Uniform Convergence Rate of the Kernel Density Estimator Adaptive to Intrinsic Volume Dimension

<http://proceedings.mlr.press/v97/kim19e.html>

AUTHORS: Jisu Kim, Jaehyeok Shin, Alessandro Rinaldo, Larry Wasserman

HIGHLIGHT: We derive concentration inequalities for the supremum norm of the difference between a kernel density estimator (KDE) and its point-wise expectation that hold uniformly over the selection of the bandwidth and under weaker conditions on the kernel and the data generating distribution than previously used in the literature.

344, TITLE: Bit-Swap: Recursive Bits-Back Coding for Lossless Compression with Hierarchical Latent Variables

<http://proceedings.mlr.press/v97/kingma19a.html>

AUTHORS: Friso Kingma, Pieter Abbeel, Jonathan Ho

HIGHLIGHT: In this paper we propose Bit-Swap, a new compression scheme that generalizes BB-ANS and achieves strictly better compression rates for hierarchical latent variable models with Markov chain structure.

345, TITLE: CompILE: Compositional Imitation Learning and Execution

<http://proceedings.mlr.press/v97/kipf19a.html>

AUTHORS: Thomas Kipf, Yujia Li, Hanjun Dai, Vinicius Zambaldi, Alvaro Sanchez-Gonzalez, Edward Grefenstette, Pushmeet Kohli, Peter Battaglia  
HIGHLIGHT: We introduce Compositional Imitation Learning and Execution (CompILE): a framework for learning reusable, variable-length segments of hierarchically-structured behavior from demonstration data.

346, TITLE: Adaptive and Safe Bayesian Optimization in High Dimensions via One-Dimensional Subspaces  
<http://proceedings.mlr.press/v97/kirschner19a.html>

AUTHORS: Johannes Kirschner, Mojmir Mutny, Nicole Hiller, Rasmus Ischebeck, Andreas Krause  
HIGHLIGHT: In order to scale the method and keep its benefits, we propose an algorithm (LineBO) that restricts the problem to a sequence of iteratively chosen one-dimensional sub-problems that can be solved efficiently.

347, TITLE: AUC: A Performance Metric for Multi-Class Machine Learning Models  
<http://proceedings.mlr.press/v97/kleiman19a.html>

AUTHORS: Ross Kleiman, David Page  
HIGHLIGHT: We provide in this work a multi-class extension of AUC that we call AUC that is derived from first principles of the binary class AUC.

348, TITLE: Fair k-Center Clustering for Data Summarization  
<http://proceedings.mlr.press/v97/kleindessner19a.html>

AUTHORS: Matthias Kleindessner, Pranjal Awasthi, Jamie Morgenstern  
HIGHLIGHT: In this paper, we resolve this gap by providing a simple approximation algorithm for the k-center problem under the fairness constraint with running time linear in the size of the data set and k.

349, TITLE: Guarantees for Spectral Clustering with Fairness Constraints  
<http://proceedings.mlr.press/v97/kleindessner19b.html>

AUTHORS: Matthias Kleindessner, Samira Samadi, Pranjal Awasthi, Jamie Morgenstern  
HIGHLIGHT: Given the widespread popularity of spectral clustering (SC) for partitioning graph data, we study a version of constrained SC in which we try to incorporate the fairness notion proposed by Chierichetti et al. (2017).

350, TITLE: POPQORN: Quantifying Robustness of Recurrent Neural Networks  
<http://proceedings.mlr.press/v97/ko19a.html>

AUTHORS: Ching-Yun Ko, Zhaoyang Lyu, Lily Weng, Luca Daniel, Ngai Wong, Dahua Lin  
HIGHLIGHT: In this work, we propose POPQORN (Propagated-output Quantified Robustness for RNNs), a general algorithm to quantify robustness of RNNs, including vanilla RNNs, LSTMs, and GRUs.

351, TITLE: Decentralized Stochastic Optimization and Gossip Algorithms with Compressed Communication  
<http://proceedings.mlr.press/v97/koloskova19a.html>

AUTHORS: Anastasia Koloskova, Sebastian Stich, Martin Jaggi  
HIGHLIGHT: We (ii) present a novel gossip algorithm, CHOCO-GOSSIP, for the average consensus problem that converges in time  $O(1/(\rho^2\delta) \log(1/\epsilon))$  for accuracy  $\epsilon > 0$ .

352, TITLE: Robust Learning from Untrusted Sources  
<http://proceedings.mlr.press/v97/konstantinov19a.html>

AUTHORS: Nikola Konstantinov, Christoph Lampert  
HIGHLIGHT: In this work, we address the question of how to learn robustly in such scenarios.

353, TITLE: Stochastic Beams and Where To Find Them: The Gumbel-Top-k Trick for Sampling Sequences Without Replacement

<http://proceedings.mlr.press/v97/kool19a.html>

AUTHORS: Wouter Kool, Herke Van Hoof, Max Welling  
HIGHLIGHT: We show how to implicitly apply this 'Gumbel-Top-k' trick on a factorized distribution over sequences, allowing to draw exact samples without replacement using a Stochastic Beam Search.

- 354, TITLE: LIT: Learned Intermediate Representation Training for Model Compression  
<http://proceedings.mlr.press/v97/koratana19a.html>  
AUTHORS: Animesh Koratana, Daniel Kang, Peter Bailis, Matei Zaharia  
HIGHLIGHT: In this work, we introduce Learned Intermediate representation Training (LIT), a novel model compression technique that outperforms a range of recent model compression techniques by leveraging the highly repetitive structure of modern DNNs (e.g., ResNet).
- 355, TITLE: Similarity of Neural Network Representations Revisited  
<http://proceedings.mlr.press/v97/kornblith19a.html>  
AUTHORS: Simon Kornblith, Mohammad Norouzi, Honglak Lee, Geoffrey Hinton  
HIGHLIGHT: We introduce a similarity index that measures the relationship between representational similarity matrices and does not suffer from this limitation.
- 356, TITLE: On the Complexity of Approximating Wasserstein Barycenters  
<http://proceedings.mlr.press/v97/kroshnin19a.html>  
AUTHORS: Alexey Kroshnin, Nazarii Tupitsa, Darina Dvinskikh, Pavel Dvurechensky, Alexander Gasnikov, Cesar Uribe  
HIGHLIGHT: To overcome this issue, we propose a novel proximal-IBP algorithm, which can be seen as a proximal gradient method, which uses IBP on each iteration to make a proximal step.
- 357, TITLE: Estimate Sequences for Variance-Reduced Stochastic Composite Optimization  
<http://proceedings.mlr.press/v97/kulunchakov19a.html>  
AUTHORS: Andrei Kulunchakov, Julien Mairal  
HIGHLIGHT: In this paper, we propose a unified view of gradient-based algorithms for stochastic convex composite optimization by extending the concept of estimate sequence introduced by Nesterov.
- 358, TITLE: Faster Algorithms for Binary Matrix Factorization  
<http://proceedings.mlr.press/v97/kumar19a.html>  
AUTHORS: Ravi Kumar, Rina Panigrahy, Ali Rahimi, David Woodruff  
HIGHLIGHT: We give faster approximation algorithms for well-studied variants of Binary Matrix Factorization (BMF), where we are given a binary  $m \times n$  matrix  $A$  and would like to find binary rank- $k$  matrices  $U, V$  to minimize the Frobenius norm of  $U \cdot V - A$ .
- 359, TITLE: Loss Landscapes of Regularized Linear Autoencoders  
<http://proceedings.mlr.press/v97/kunin19a.html>  
AUTHORS: Daniel Kunin, Jonathan Bloom, Aleksandrina Goeva, Cotton Seed  
HIGHLIGHT: In this paper, we prove that  $L_2$ -regularized LAEs are symmetric at all critical points and learn the principal directions as the left singular vectors of the decoder.
- 360, TITLE: Geometry and Symmetry in Short-and-Sparse Deconvolution  
<http://proceedings.mlr.press/v97/kuo19a.html>  
AUTHORS: Han-Wen Kuo, Yenson Lau, Yuqian Zhang, John Wright  
HIGHLIGHT: We propose a method based on nonconvex optimization, which under certain conditions recovers the target short and sparse signals, up to a signed shift symmetry which is intrinsic to this model.
- 361, TITLE: A Large-Scale Study on Regularization and Normalization in GANs  
<http://proceedings.mlr.press/v97/kurach19a.html>  
AUTHORS: Karol Kurach, Mario Lucic, Xiaohua Zhai, Marcin Michalski, Sylvain Gelly  
HIGHLIGHT: In this work we take a sober view of the current state of GANs from a practical perspective.
- 362, TITLE: Making Decisions that Reduce Discriminatory Impacts  
<http://proceedings.mlr.press/v97/kusner19a.html>  
AUTHORS: Matt Kusner, Chris Russell, Joshua Loftus, Ricardo Silva

**HIGHLIGHT:** To address this, we describe causal methods that model the relevant parts of the real-world system in which the decisions are made.

**363, TITLE:** Garbage In, Reward Out: Bootstrapping Exploration in Multi-Armed Bandits  
<http://proceedings.mlr.press/v97/kveton19a.html>  
**AUTHORS:** Branislav Kveton, Csaba Szepesvari, Sharan Vaswani, Zheng Wen, Tor Lattimore, Mohammad Ghavamzadeh  
**HIGHLIGHT:** We propose a bandit algorithm that explores by randomizing its history of rewards.

**364, TITLE:** Characterizing Well-Behaved vs. Pathological Deep Neural Networks  
<http://proceedings.mlr.press/v97/labatie19a.html>  
**AUTHORS:** Antoine Labatie  
**HIGHLIGHT:** We introduce a novel approach, requiring only mild assumptions, for the characterization of deep neural networks at initialization.

**365, TITLE:** State-Reification Networks: Improving Generalization by Modeling the Distribution of Hidden Representations  
<http://proceedings.mlr.press/v97/lamb19a.html>  
**AUTHORS:** Alex Lamb, Jonathan Binas, Anirudh Goyal, Sandeep Subramanian, Ioannis Mitliagkas, Yoshua Bengio, Michael Mozer  
**HIGHLIGHT:** We introduce a method, which we refer to as `_state reification_`, that involves modeling the distribution of hidden states over the training data and then projecting hidden states observed during testing toward this distribution.

**366, TITLE:** A Recurrent Neural Cascade-based Model for Continuous-Time Diffusion  
<http://proceedings.mlr.press/v97/lamprier19a.html>  
**AUTHORS:** Sylvain Lamprier  
**HIGHLIGHT:** In this paper we propose a model at the crossroads of these two extremes, which embeds the history of diffusion in infected nodes as hidden continuous states.

**367, TITLE:** Projection onto Minkowski Sums with Application to Constrained Learning  
<http://proceedings.mlr.press/v97/lange19a.html>  
**AUTHORS:** Kenneth Lange, Joong-Ho Won, Jason Xu  
**HIGHLIGHT:** We introduce block descent algorithms for projecting onto Minkowski sums of sets.

**368, TITLE:** Safe Policy Improvement with Baseline Bootstrapping  
<http://proceedings.mlr.press/v97/laroche19a.html>  
**AUTHORS:** Romain Laroche, Paul Trichelair, Remi Tachet Des Combes  
**HIGHLIGHT:** This paper considers Safe Policy Improvement (SPI) in Batch Reinforcement Learning (Batch RL): from a fixed dataset and without direct access to the true environment, train a policy that is guaranteed to perform at least as well as the baseline policy used to collect the data.

**369, TITLE:** A Better k-means++ Algorithm via Local Search  
<http://proceedings.mlr.press/v97/lattanzi19a.html>  
**AUTHORS:** Silvio Lattanzi, Christian Sohler  
**HIGHLIGHT:** In this paper, we develop a new variant of k-means++ seeding that in expectation achieves a constant approximation guarantee.

**370, TITLE:** Lorentzian Distance Learning for Hyperbolic Representations  
<http://proceedings.mlr.press/v97/law19a.html>  
**AUTHORS:** Marc Law, Renjie Liao, Jake Snell, Richard Zemel  
**HIGHLIGHT:** We introduce an approach to learn representations based on the Lorentzian distance in hyperbolic geometry.

- 371, TITLE: DP-GP-LVM: A Bayesian Non-Parametric Model for Learning Multivariate Dependency Structures  
http://proceedings.mlr.press/v97/lawrence19a.html  
AUTHORS: Andrew Lawrence, Carl Henrik Ek, Neill Campbell  
HIGHLIGHT: We present a non-parametric Bayesian latent variable model capable of learning dependency structures across dimensions in a multivariate setting.
- 372, TITLE: POLITEX: Regret Bounds for Policy Iteration using Expert Prediction  
http://proceedings.mlr.press/v97/lazic19a.html  
AUTHORS: Nevena Lazic, Yasin Abbasi-Yadkori, Kush Bhatia, Gellert Weisz, Peter Bartlett, Csaba Szepesvari  
HIGHLIGHT: We present POLITEX (POLicy Iteration with EXpert advice), a variant of policy iteration where each policy is a Boltzmann distribution over the sum of action-value function estimates of the previous policies, and analyze its regret in continuing RL problems.
- 373, TITLE: Batch Policy Learning under Constraints  
http://proceedings.mlr.press/v97/le19a.html  
AUTHORS: Hoang Le, Cameron Voloshin, Yisong Yue  
HIGHLIGHT: As part of off-policy learning, we propose a simple method for off-policy policy evaluation (OPE) and derive PAC-style bounds.
- 374, TITLE: Target-Based Temporal-Difference Learning  
http://proceedings.mlr.press/v97/lee19a.html  
AUTHORS: Donghwan Lee, Niao He  
HIGHLIGHT: In this work, we introduce a new family of target-based temporal difference (TD) learning algorithms that maintain two separate learning parameters  $\{-\}$  the target variable and online variable.
- 375, TITLE: Functional Transparency for Structured Data: a Game-Theoretic Approach  
http://proceedings.mlr.press/v97/lee19b.html  
AUTHORS: Guang-He Lee, Wengong Jin, David Alvarez-Melis, Tommi Jaakkola  
HIGHLIGHT: We provide a new approach to training neural models to exhibit transparency in a well-defined, functional manner.
- 376, TITLE: Self-Attention Graph Pooling  
http://proceedings.mlr.press/v97/lee19c.html  
AUTHORS: Junhyun Lee, Inyeop Lee, Jaewoo Kang  
HIGHLIGHT: In this paper, we propose a graph pooling method based on self-attention.
- 377, TITLE: Set Transformer: A Framework for Attention-based Permutation-Invariant Neural Networks  
http://proceedings.mlr.press/v97/lee19d.html  
AUTHORS: Juho Lee, Yoonho Lee, Jungtaek Kim, Adam Kosior, Seungjin Choi, Yee Whye Teh  
HIGHLIGHT: We present an attention-based neural network module, the Set Transformer, specifically designed to model interactions among elements in the input set.
- 378, TITLE: First-Order Algorithms Converge Faster than  $\mathcal{O}(1/k)$  on Convex Problems  
http://proceedings.mlr.press/v97/lee19e.html  
AUTHORS: Ching-Pei Lee, Stephen Wright  
HIGHLIGHT: In this work, we improve this rate to  $\mathcal{o}(1/k)$ .
- 379, TITLE: Robust Inference via Generative Classifiers for Handling Noisy Labels  
http://proceedings.mlr.press/v97/lee19f.html  
AUTHORS: Kimin Lee, Sukmin Yun, Kibok Lee, Honglak Lee, Bo Li, Jinwoo Shin  
HIGHLIGHT: To mitigate the issue, we propose a novel inference method, termed Robust Generative classifier (RoG), applicable to any discriminative (e.g., softmax) neural classifier pre-trained on noisy datasets.

- 380, TITLE: Sublinear Time Nearest Neighbor Search over Generalized Weighted Space  
<http://proceedings.mlr.press/v97/lei19a.html>  
AUTHORS: Yifan Lei, Qiang Huang, Mohan Kankanhalli, Anthony Tung  
HIGHLIGHT: Based on the idea of Asymmetric Locality-Sensitive Hashing (ALSH), we introduce a novel spherical asymmetric transformation and propose the first two novel weight-oblivious hashing schemes SL-ALSH and S2-ALSH accordingly.
- 381, TITLE: MONK Outlier-Robust Mean Embedding Estimation by Median-of-Means  
<http://proceedings.mlr.press/v97/lerasle19a.html>  
AUTHORS: Matthieu Lerasle, Zoltan Szabo, Timothee Mathieu, Guillaume Lecue  
HIGHLIGHT: In this paper, we show how the recently emerged principle of median-of-means can be used to design estimators for kernel mean embedding and MMD with excessive resistance properties to outliers, and optimal sub-Gaussian deviation bounds under mild assumptions.
- 382, TITLE: Cheap Orthogonal Constraints in Neural Networks: A Simple Parametrization of the Orthogonal and Unitary Group  
<http://proceedings.mlr.press/v97/lezciano-casado19a.html>  
AUTHORS: Mario Lezciano-Casado, David Martinez-Rubio  
HIGHLIGHT: We introduce a novel approach to perform first-order optimization with orthogonal and unitary constraints.
- 383, TITLE: Are Generative Classifiers More Robust to Adversarial Attacks?  
<http://proceedings.mlr.press/v97/li19a.html>  
AUTHORS: Yingzhen Li, John Bradshaw, Yash Sharma  
HIGHLIGHT: In this paper, we propose and investigate the deep Bayes classifier, which improves classical naive Bayes with conditional deep generative models.
- 384, TITLE: Sublinear quantum algorithms for training linear and kernel-based classifiers  
<http://proceedings.mlr.press/v97/li19b.html>  
AUTHORS: Tongyang Li, Shouvanik Chakrabarti, Xiaodi Wu  
HIGHLIGHT: We investigate quantum algorithms for classification, a fundamental problem in machine learning, with provable guarantees.
- 385, TITLE: LGM-Net: Learning to Generate Matching Networks for Few-Shot Learning  
<http://proceedings.mlr.press/v97/li19c.html>  
AUTHORS: Huaiyu Li, Weiming Dong, Xing Mei, Chongyang Ma, Feiyue Huang, Bao-Gang Hu  
HIGHLIGHT: In this work, we propose a novel meta-learning approach for few-shot classification, which learns transferable prior knowledge across tasks and directly produces network parameters for similar unseen tasks with training samples.
- 386, TITLE: Graph Matching Networks for Learning the Similarity of Graph Structured Objects  
<http://proceedings.mlr.press/v97/li19d.html>  
AUTHORS: Yujia Li, Chenjie Gu, Thomas Dullien, Oriol Vinyals, Pushmeet Kohli  
HIGHLIGHT: This paper addresses the challenging problem of retrieval and matching of graph structured objects, and makes two key contributions.
- 387, TITLE: Area Attention  
<http://proceedings.mlr.press/v97/li19e.html>  
AUTHORS: Yang Li, Lukasz Kaiser, Samy Bengio, Si Si  
HIGHLIGHT: We propose area attention: a way to attend to areas in the memory, where each area contains a group of items that are structurally adjacent, e.g., spatially for a 2D memory such as images, or temporally for a 1D memory such as natural language sentences.

- 388, TITLE: Online Learning to Rank with Features  
<http://proceedings.mlr.press/v97/li19f.html>  
AUTHORS: Shuai Li, Tor Lattimore, Csaba Szepesvari  
HIGHLIGHT: We introduce a new model for online ranking in which the click probability factors into an examination and attractiveness function and the attractiveness function is a linear function of a feature vector and an unknown parameter.
- 389, TITLE: NATTACK: Learning the Distributions of Adversarial Examples for an Improved Black-Box Attack on Deep Neural Networks  
<http://proceedings.mlr.press/v97/li19g.html>  
AUTHORS: Yandong Li, Lijun Li, Liqiang Wang, Tong Zhang, Boqing Gong  
HIGHLIGHT: In this paper, we propose a black-box adversarial attack algorithm that can defeat both vanilla DNNs and those generated by various defense techniques developed recently.
- 390, TITLE: Bayesian Joint Spike-and-Slab Graphical Lasso  
<http://proceedings.mlr.press/v97/li19h.html>  
AUTHORS: Zehang Li, Tyler McCormick, Samuel Clark  
HIGHLIGHT: In this article, we propose a new class of priors for Bayesian inference with multiple Gaussian graphical models.
- 391, TITLE: Exploiting Worker Correlation for Label Aggregation in Crowdsourcing  
<http://proceedings.mlr.press/v97/li19i.html>  
AUTHORS: Yuan Li, Benjamin Rubinstein, Trevor Cohn  
HIGHLIGHT: In this paper, we argue that existing crowdsourcing approaches do not sufficiently model worker correlations observed in practical settings; we propose in response an enhanced Bayesian classifier combination (EBCC) model, with inference based on a mean-field variational approach.
- 392, TITLE: Adversarial camera stickers: A physical camera-based attack on deep learning systems  
<http://proceedings.mlr.press/v97/li19j.html>  
AUTHORS: Juncheng Li, Frank Schmidt, Zico Kolter  
HIGHLIGHT: In this work, we consider an alternative question: is it possible to fool deep classifiers, over all perceived objects of a certain type, by physically manipulating the camera itself?
- 393, TITLE: Towards a Unified Analysis of Random Fourier Features  
<http://proceedings.mlr.press/v97/li19k.html>  
AUTHORS: Zhu Li, Jean-Francois Ton, Dino Oglic, Dino Sejdinovic  
HIGHLIGHT: We study both the standard random Fourier features method for which we improve the existing bounds on the number of features required to guarantee the corresponding minimax risk convergence rate of kernel ridge regression, as well as a data-dependent modification which samples features proportional to ridge leverage scores and further reduces the required number of features.
- 394, TITLE: Feature-Critic Networks for Heterogeneous Domain Generalization  
<http://proceedings.mlr.press/v97/li19l.html>  
AUTHORS: Yiying Li, Yongxin Yang, Wei Zhou, Timothy Hospedales  
HIGHLIGHT: In this work, we propose a learning to learn approach, where the auxiliary loss that helps generalisation is itself learned.
- 395, TITLE: Learn to Grow: A Continual Structure Learning Framework for Overcoming Catastrophic Forgetting  
<http://proceedings.mlr.press/v97/li19m.html>  
AUTHORS: Xilai Li, Yingbo Zhou, Tianfu Wu, Richard Socher, Caiming Xiong  
HIGHLIGHT: This paper presents a conceptually simple yet general and effective framework for handling catastrophic forgetting in continual learning with DNNs.

- 396, TITLE: Alternating Minimizations Converge to Second-Order Optimal Solutions  
<http://proceedings.mlr.press/v97/li19n.html>  
AUTHORS: Qiuwei Li, Zhihui Zhu, Gongguo Tang  
HIGHLIGHT: We show that under mild assumptions on the (nonconvex) objective function, both algorithms avoid strict saddles almost surely from random initialization.
- 397, TITLE: Cautious Regret Minimization: Online Optimization with Long-Term Budget Constraints  
<http://proceedings.mlr.press/v97/liakopoulos19a.html>  
AUTHORS: Nikolaos Liakopoulos, Apostolos Destounis, Georgios Paschos, Thrasyvoulos Spyropoulos, Panayotis Mertikopoulos  
HIGHLIGHT: We study a class of online convex optimization problems with long-term budget constraints that arise naturally as reliability guarantees or total consumption constraints.
- 398, TITLE: Regularization in directable environments with application to Tetris  
<http://proceedings.mlr.press/v97/lichtenberg19a.html>  
AUTHORS: Jan Lichtenberg, Simsek  
HIGHLIGHT: We present a regularized linear model called STEW that benefits from a generic and prevalent form of prior knowledge: feature directions.
- 399, TITLE: Inference and Sampling of  $\mathbb{K}_{33}$ -free Ising Models  
<http://proceedings.mlr.press/v97/likhosherstov19a.html>  
AUTHORS: Valerii Likhosherstov, Yury Maximov, Misha Chertkov  
HIGHLIGHT: Inference and Sampling of  $\mathbb{K}_{33}$ -free Ising Models.
- 400, TITLE: Kernel-Based Reinforcement Learning in Robust Markov Decision Processes  
<http://proceedings.mlr.press/v97/lim19a.html>  
AUTHORS: Shiau Hong Lim, Arnaud Autef  
HIGHLIGHT: We extend these results to the much larger class of kernel-based approximators and show, both analytically and empirically that the robust policies can significantly outperform the non-robust counterpart.
- 401, TITLE: On Efficient Optimal Transport: An Analysis of Greedy and Accelerated Mirror Descent Algorithms  
<http://proceedings.mlr.press/v97/lin19a.html>  
AUTHORS: Tianyi Lin, Nhat Ho, Michael Jordan  
HIGHLIGHT: We provide theoretical analyses for two algorithms that solve the regularized optimal transport (OT) problem between two discrete probability measures with at most  $n$  atoms.
- 402, TITLE: Fast and Simple Natural-Gradient Variational Inference with Mixture of Exponential-family Approximations  
<http://proceedings.mlr.press/v97/lin19b.html>  
AUTHORS: Wu Lin, Mohammad Emtiyaz Khan, Mark Schmidt  
HIGHLIGHT: In this paper, we extend their application to estimate structured approximations such as mixtures of EF distributions.
- 403, TITLE: Acceleration of SVRG and Katyusha X by Inexact Preconditioning  
<http://proceedings.mlr.press/v97/liu19a.html>  
AUTHORS: Yanli Liu, Fei Feng, Wotao Yin  
HIGHLIGHT: In this paper, we propose to accelerate these two algorithms by inexact preconditioning, the proposed methods employ fixed preconditioners, although the subproblem in each epoch becomes harder, it suffices to apply fixed number of simple subroutines to solve it inexactly, without losing the overall convergence.
- 404, TITLE: Transferable Adversarial Training: A General Approach to Adapting Deep Classifiers  
<http://proceedings.mlr.press/v97/liu19b.html>



AUTHORS: Hong Liu, Mingsheng Long, Jianmin Wang, Michael Jordan  
HIGHLIGHT: To this end, we propose Transferable Adversarial Training (TAT) to enable the adaptation of deep classifiers.

405, TITLE: Rao-Blackwellized Stochastic Gradients for Discrete Distributions  
<http://proceedings.mlr.press/v97/liu19c.html>  
AUTHORS: Runjing Liu, Jeffrey Regier, Nilesh Tripuraneni, Michael Jordan, Jon McAuliffe  
HIGHLIGHT: In this paper, we describe a technique that can be applied to reduce the variance of any such estimator, without changing its bias {—} in particular, unbiasedness is retained.

406, TITLE: Sparse Extreme Multi-label Learning with Oracle Property  
<http://proceedings.mlr.press/v97/liu19d.html>  
AUTHORS: Weiwei Liu, Xiaobo Shen  
HIGHLIGHT: To fill this gap, we present a unified framework for SLEEC with nonconvex penalty.

407, TITLE: Data Poisoning Attacks on Stochastic Bandits  
<http://proceedings.mlr.press/v97/liu19e.html>  
AUTHORS: Fang Liu, Ness Shroff  
HIGHLIGHT: In this paper, we propose a framework of offline attacks on bandit algorithms and study convex optimization based attacks on several popular bandit algorithms.

408, TITLE: The Implicit Fairness Criterion of Unconstrained Learning  
<http://proceedings.mlr.press/v97/liu19f.html>  
AUTHORS: Lydia T. Liu, Max Simchowitz, Moritz Hardt  
HIGHLIGHT: We clarify what fairness guarantees we can and cannot expect to follow from unconstrained machine learning.

409, TITLE: Taming MAML: Efficient unbiased meta-reinforcement learning  
<http://proceedings.mlr.press/v97/liu19g.html>  
AUTHORS: Hao Liu, Richard Socher, Caiming Xiong  
HIGHLIGHT: We propose a surrogate objective function named, Taming MAML (TMAML), that adds control variates into gradient estimation via automatic differentiation.

410, TITLE: On Certifying Non-Uniform Bounds against Adversarial Attacks  
<http://proceedings.mlr.press/v97/liu19h.html>  
AUTHORS: Chen Liu, Ryota Tomioka, Volkan Cevher  
HIGHLIGHT: We formulate our target as an optimization problem with nonlinear constraints.

411, TITLE: Understanding and Accelerating Particle-Based Variational Inference  
<http://proceedings.mlr.press/v97/liu19i.html>  
AUTHORS: Chang Liu, Jingwei Zhuo, Pengyu Cheng, Ruiyi Zhang, Jun Zhu  
HIGHLIGHT: We propose an acceleration framework and a principled bandwidth-selection method for general ParVIs; these are based on the developed theory and leverage the geometry of the Wasserstein space.

412, TITLE: Understanding MCMC Dynamics as Flows on the Wasserstein Space  
<http://proceedings.mlr.press/v97/liu19j.html>  
AUTHORS: Chang Liu, Jingwei Zhuo, Jun Zhu  
HIGHLIGHT: In this work, by developing novel concepts, we propose a theoretical framework that recognizes a general MCMC dynamics as the fiber-gradient Hamiltonian flow on the Wasserstein space of a fiber-Riemannian Poisson manifold.

413, TITLE: Sliced-Wasserstein Flows: Nonparametric Generative Modeling via Optimal Transport and Diffusions

<http://proceedings.mlr.press/v97/liutkus19a.html>

AUTHORS: Antoine Liutkus, Umut Simsekli, Szymon Majewski, Alain Durmus, Fabian-Robert St?ter  
HIGHLIGHT: By building upon the recent theory that established the connection between implicit generative modeling (IGM) and optimal transport, in this study, we propose a novel parameter-free algorithm for learning the underlying distributions of complicated datasets and sampling from them.

414, TITLE: Challenging Common Assumptions in the Unsupervised Learning of Disentangled Representations  
<http://proceedings.mlr.press/v97/locatello19a.html>

AUTHORS: Francesco Locatello, Stefan Bauer, Mario Lucic, Gunnar Raetsch, Sylvain Gelly, Bernhard Sch?lkopf, Olivier Bachem  
HIGHLIGHT: In this paper, we provide a sober look at recent progress in the field and challenge some common assumptions.

415, TITLE: Bayesian Counterfactual Risk Minimization

<http://proceedings.mlr.press/v97/london19a.html>

AUTHORS: Ben London, Ted Sandler  
HIGHLIGHT: We present a Bayesian view of counterfactual risk minimization (CRM) for offline learning from logged bandit feedback.

416, TITLE: PA-GD: On the Convergence of Perturbed Alternating Gradient Descent to Second-Order Stationary Points for Structured Nonconvex Optimization

<http://proceedings.mlr.press/v97/lu19a.html>

AUTHORS: Songtao Lu, Mingyi Hong, Zhengdao Wang  
HIGHLIGHT: In this paper, we consider a smooth unconstrained nonconvex optimization problem, and propose a perturbed A-GD (PA-GD) which is able to converge (with high probability) to the second-order stationary points (SOSPs) with a global sublinear rate.

417, TITLE: Neurally-Guided Structure Inference

<http://proceedings.mlr.press/v97/lu19b.html>

AUTHORS: Sidi Lu, Jiayuan Mao, Joshua Tenenbaum, Jiajun Wu  
HIGHLIGHT: In this paper, we propose a hybrid inference algorithm, the Neurally-Guided Structure Inference (NG-SI), keeping the advantages of both search-based and data-driven methods.

418, TITLE: Optimal Algorithms for Lipschitz Bandits with Heavy-tailed Rewards

<http://proceedings.mlr.press/v97/lu19c.html>

AUTHORS: Shiyin Lu, Guanghui Wang, Yao Hu, Lijun Zhang  
HIGHLIGHT: To address this limitation, in this paper we relax the assumption on rewards to allow arbitrary distributions that have finite  $(1+\epsilon)$ -th moments for some  $\epsilon \in (0, 1]$ , and propose algorithms that enjoy a sublinear regret of  $\tilde{O}(T^{\frac{1}{d_z \epsilon + 1}})$  where  $T$  is the time horizon and  $d_z$  is the zooming dimension.

419, TITLE: CoT: Cooperative Training for Generative Modeling of Discrete Data

<http://proceedings.mlr.press/v97/lu19d.html>

AUTHORS: Sidi Lu, Lantao Yu, Siyuan Feng, Yaoming Zhu, Weinan Zhang  
HIGHLIGHT: In this paper, we study the generative models of sequential discrete data.

420, TITLE: Generalized Approximate Survey Propagation for High-Dimensional Estimation

<http://proceedings.mlr.press/v97/lucibello19a.html>

AUTHORS: Carlo Lucibello, Luca Saglietti, Yue Lu  
HIGHLIGHT: In this paper, we propose a new algorithm, named Generalized Approximate Survey Propagation (GASP), for solving GLE in the presence of prior or model misspecifications. Furthermore, we present a set of state evolution equations that can precisely characterize the performance of GASP in the high-dimensional limit.

- 421, TITLE: High-Fidelity Image Generation With Fewer Labels  
<http://proceedings.mlr.press/v97/lucic19a.html>  
AUTHORS: Mario Lucic, Michael Tschannen, Marvin Ritter, Xiaohua Zhai, Olivier Bachem, Sylvain Gelly  
HIGHLIGHT: In this work we demonstrate how one can benefit from recent work on self- and semi-supervised learning to outperform the state of the art on both unsupervised ImageNet synthesis, as well as in the conditional setting.
- 422, TITLE: Leveraging Low-Rank Relations Between Surrogate Tasks in Structured Prediction  
<http://proceedings.mlr.press/v97/luise19a.html>  
AUTHORS: Giulia Luise, Dimitrios Stamos, Massimiliano Pontil, Carlo Ciliberto  
HIGHLIGHT: We propose an efficient algorithm based on trace norm regularization which, differently from previous methods, does not require explicit knowledge of the coding/decoding functions of the surrogate framework.
- 423, TITLE: Differentiable Dynamic Normalization for Learning Deep Representation  
<http://proceedings.mlr.press/v97/luo19a.html>  
AUTHORS: Ping Luo, Peng Zhanglin, Shao Wenqi, Zhang Ruimao, Ren Jiamin, Wu Lingyun  
HIGHLIGHT: This work presents Dynamic Normalization (DN), which is able to learn arbitrary normalization operations for different convolutional layers in a deep ConvNet.
- 424, TITLE: Disentangled Graph Convolutional Networks  
<http://proceedings.mlr.press/v97/ma19a.html>  
AUTHORS: Jianxin Ma, Peng Cui, Kun Kuang, Xin Wang, Wenwu Zhu  
HIGHLIGHT: In this paper, we introduce the disentangled graph convolutional network (DisenGCN) to learn disentangled node representations.
- 425, TITLE: Variational Implicit Processes  
<http://proceedings.mlr.press/v97/ma19b.html>  
AUTHORS: Chao Ma, Yingzhen Li, Jose Miguel Hernandez-Lobato  
HIGHLIGHT: We introduce the implicit processes (IPs), a stochastic process that places implicitly defined multivariate distributions over any finite collections of random variables.
- 426, TITLE: EDDI: Efficient Dynamic Discovery of High-Value Information with Partial VAE  
<http://proceedings.mlr.press/v97/ma19c.html>  
AUTHORS: Chao Ma, Sebastian Tschiatschek, Konstantina Palla, Jose Miguel Hernandez-Lobato, Sebastian Nowozin, Cheng Zhang  
HIGHLIGHT: To this end, we propose a principled framework, named EDDI (Efficient Dynamic Discovery of high-value Information), based on the theory of Bayesian experimental design.
- 427, TITLE: Bayesian leave-one-out cross-validation for large data  
<http://proceedings.mlr.press/v97/magnusson19a.html>  
AUTHORS: M?ns Magnusson, Michael Andersen, Johan Jonasson, Aki Vehtari  
HIGHLIGHT: We propose a combination of using approximate inference techniques and probability-proportional-to-size-sampling (PPS) for fast LOO model evaluation for large datasets.
- 428, TITLE: Composable Core-sets for Determinant Maximization: A Simple Near-Optimal Algorithm  
<http://proceedings.mlr.press/v97/mahabadi19a.html>  
AUTHORS: Sepideh Mahabadi, Piotr Indyk, Shayan Oveis Gharan, Alireza Rezaei  
HIGHLIGHT: In this work, we consider efficient construction of composable core-sets for the determinant maximization problem.
- 429, TITLE: Guided evolutionary strategies: augmenting random search with surrogate gradients  
<http://proceedings.mlr.press/v97/maheswaranathan19a.html>  
AUTHORS: Niru Maheswaranathan, Luke Metz, George Tucker, Dami Choi, Jascha Sohl-Dickstein

**HIGHLIGHT:** We propose Guided Evolutionary Strategies (GES), a method for optimally using surrogate gradient directions to accelerate random search.

430, **TITLE:** Data Poisoning Attacks in Multi-Party Learning

<http://proceedings.mlr.press/v97/mahloujifar19a.html>

**AUTHORS:** Saeed Mahloujifar, Mohammad Mahmoody, Ameer Mohammed

**HIGHLIGHT:** In this work, we demonstrate universal multi-party poisoning attacks that adapt and apply to any multi-party learning process with arbitrary interaction pattern between the parties.

431, **TITLE:** Traditional and Heavy Tailed Self Regularization in Neural Network Models

<http://proceedings.mlr.press/v97/mahoney19a.html>

**AUTHORS:** Michael Mahoney, Charles Martin

**HIGHLIGHT:** Building on recent results in RMT, most notably its extension to Universality classes of Heavy-Tailed matrices, we develop a theory to identify 5+1 Phases of Training, corresponding to increasing amounts of Implicit Self-Regularization.

432, **TITLE:** Curvature-Exploiting Acceleration of Elastic Net Computations

<http://proceedings.mlr.press/v97/mai19a.html>

**AUTHORS:** Vien Mai, Mikael Johansson

**HIGHLIGHT:** This paper introduces an efficient second-order method for solving the elastic net problem.

433, **TITLE:** Breaking the gridlock in Mixture-of-Experts: Consistent and Efficient Algorithms

<http://proceedings.mlr.press/v97/makkuva19a.html>

**AUTHORS:** Ashok Makkuva, Pramod Viswanath, Sreeram Kannan, Sewoong Oh

**HIGHLIGHT:** In this paper, we introduce the first algorithm that learns the true parameters of a MoE model for a wide class of non-linearities with global consistency guarantees.

434, **TITLE:** Calibrated Model-Based Deep Reinforcement Learning

<http://proceedings.mlr.press/v97/malik19a.html>

**AUTHORS:** Ali Malik, Volodymyr Kuleshov, Jiaming Song, Danny Nemer, Harlan Seymour, Stefano Ermon

**HIGHLIGHT:** We describe a simple way to augment any model-based reinforcement learning agent with a calibrated model and show that doing so consistently improves planning, sample complexity, and exploration.

435, **TITLE:** Learning from Delayed Outcomes via Proxies with Applications to Recommender Systems

<http://proceedings.mlr.press/v97/mann19a.html>

**AUTHORS:** Timothy Arthur Mann, Sven Gowal, Andras Gyorgy, Huiyi Hu, Ray Jiang, Balaji Lakshminarayanan, Prav Srinivasan

**HIGHLIGHT:** Motivated by our regret analysis, we propose two neural network architectures: Factored Forecaster (FF) which is ideal if the proxy is informative of the outcome in hindsight, and Residual Factored Forecaster (RFF) that is robust to a non-informative proxy.

436, **TITLE:** Passed & Spurious: Descent Algorithms and Local Minima in Spiked Matrix-Tensor Models

<http://proceedings.mlr.press/v97/mannelli19a.html>

**AUTHORS:** Stefano Sarao Mannelli, Florent Krzakala, Pierfrancesco Urbani, Lenka Zdeborova

**HIGHLIGHT:** In this work we analyse quantitatively the interplay between the loss landscape and performance of descent algorithms in a prototypical inference problem, the spiked matrix-tensor model.

437, **TITLE:** A Baseline for Any Order Gradient Estimation in Stochastic Computation Graphs

<http://proceedings.mlr.press/v97/mao19a.html>

**AUTHORS:** Jingkai Mao, Jakob Foerster, Tim Rocktäschel, Maruan Al-Shedivat, Gregory Farquhar, Shimon Whiteson

**HIGHLIGHT:** To improve the sample efficiency of DiCE, we propose a new baseline term for higher order gradient estimation.

- 438, TITLE: Adversarial Generation of Time-Frequency Features with application in audio synthesis  
<http://proceedings.mlr.press/v97/marafiotti19a.html>  
AUTHORS: Andr?s Marafioti, Nathana?l Perraudin, Nicki Holighaus, Piotr Majdak  
HIGHLIGHT: In this article, focusing on the short-time Fourier transform, we discuss the challenges that arise in audio synthesis based on generated invertible TF features and how to overcome them.
- 439, TITLE: On the Universality of Invariant Networks  
<http://proceedings.mlr.press/v97/maron19a.html>  
AUTHORS: Haggai Maron, Ethan Fetaya, Nimrod Segol, Yaron Lipman  
HIGHLIGHT: In this paper, we consider a fundamental question that has received very little attention to date: Can these networks approximate any (continuous) invariant function?
- 440, TITLE: Decomposing feature-level variation with Covariate Gaussian Process Latent Variable Models  
<http://proceedings.mlr.press/v97/martens19a.html>  
AUTHORS: Kaspar Martens, Kieran Campbell, Christopher Yau  
HIGHLIGHT: In this paper, we propose to achieve this through a structured kernel decomposition in a hybrid Gaussian Process model which we call the Covariate Gaussian Process Latent Variable Model (c-GPLVM).
- 441, TITLE: Fairness-Aware Learning for Continuous Attributes and Treatments  
<http://proceedings.mlr.press/v97/mary19a.html>  
AUTHORS: Jeremie Mary, Cl?ment Calauz?nes, Noureddine El Karoui  
HIGHLIGHT: As common fairness metrics can be expressed as measures of (conditional) independence between variables, we propose to use the R?nyi maximum correlation coefficient to generalize fairness measurement to continuous variables.
- 442, TITLE: Optimal Minimal Margin Maximization with Boosting  
<http://proceedings.mlr.press/v97/mathiasen19a.html>  
AUTHORS: Alexander Mathiasen, Kasper Green Larsen, Allan Gr?nlund  
HIGHLIGHT: Our main contribution is a new algorithm refuting this conjecture.
- 443, TITLE: Disentangling Disentanglement in Variational Autoencoders  
<http://proceedings.mlr.press/v97/mathieu19a.html>  
AUTHORS: Emile Mathieu, Tom Rainforth, N Siddharth, Yee Whye Teh  
HIGHLIGHT: We develop a generalisation of disentanglement in variational autoencoders (VAEs)—decomposition of the latent representation—characterising it as the fulfilment of two factors: a) the latent encodings of the data having an appropriate level of overlap, and b) the aggregate encoding of the data conforming to a desired structure, represented through the prior.
- 444, TITLE: MIWAE: Deep Generative Modelling and Imputation of Incomplete Data Sets  
<http://proceedings.mlr.press/v97/mattei19a.html>  
AUTHORS: Pierre-Alexandre Mattei, Jes Frellsen  
HIGHLIGHT: We consider the problem of handling missing data with deep latent variable models (DLVMs).
- 445, TITLE: Distributional Reinforcement Learning for Efficient Exploration  
<http://proceedings.mlr.press/v97/mavrin19a.html>  
AUTHORS: Borislav Mavrin, Hengshuai Yao, Linglong Kong, Kaiwen Wu, Yaoliang Yu  
HIGHLIGHT: We propose a novel and efficient exploration method for deep RL that has two components.
- 446, TITLE: Graphical-model based estimation and inference for differential privacy  
<http://proceedings.mlr.press/v97/mckenna19a.html>  
AUTHORS: Ryan Mckenna, Daniel Sheldon, Gerome Miklau

**HIGHLIGHT:** In this work, we provide an approach to solve this estimation problem efficiently using graphical models, which is particularly effective when the distribution is high-dimensional but the measurements are over low-dimensional marginals.

447, **TITLE:** Efficient Amortised Bayesian Inference for Hierarchical and Nonlinear Dynamical Systems  
<http://proceedings.mlr.press/v97/meeds19a.html>

**AUTHORS:** Ted Meeds, Geoffrey Roeder, Paul Grant, Andrew Phillips, Neil Dalchau

**HIGHLIGHT:** We introduce a flexible, scalable Bayesian inference framework for nonlinear dynamical systems characterised by distinct and hierarchical variability at the individual, group, and population levels.

448, **TITLE:** Toward Controlling Discrimination in Online Ad Auctions  
<http://proceedings.mlr.press/v97/mehrotra19a.html>

**AUTHORS:** Anay Mehrotra, Elisa Celis, Nisheeth Vishnoi

**HIGHLIGHT:** To prevent this, we propose a constrained ad auction framework that maximizes the platform's revenue conditioned on ensuring that the audience seeing an advertiser's ad is distributed appropriately across sensitive types such as gender or race.

449, **TITLE:** Stochastic Blockmodels meet Graph Neural Networks  
<http://proceedings.mlr.press/v97/mehta19a.html>

**AUTHORS:** Nikhil Mehta, Lawrence Carin Duke, Piyush Rai

**HIGHLIGHT:** In this work, we unify these two directions by developing a sparse variational autoencoder for graphs, that retains the interpretability of SBMs, while also enjoying the excellent predictive performance of graph neural nets.

450, **TITLE:** Imputing Missing Events in Continuous-Time Event Streams  
<http://proceedings.mlr.press/v97/mei19a.html>

**AUTHORS:** Hongyuan Mei, Guanghui Qin, Jason Eisner

**HIGHLIGHT:** Given a probability model of complete sequences, we propose particle smoothing—a form of sequential importance sampling—to impute the missing events in an incomplete sequence.

451, **TITLE:** Same, Same But Different: Recovering Neural Network Quantization Error Through Weight Factorization  
<http://proceedings.mlr.press/v97/meller19a.html>

**AUTHORS:** Eldad Meller, Alexander Finkelstein, Uri Almog, Mark Grobman

**HIGHLIGHT:** In this paper, we exploit an oft-overlooked degree of freedom in most networks - for a given layer, individual output channels can be scaled by any factor provided that the corresponding weights of the next layer are inversely scaled.

452, **TITLE:** The Wasserstein Transform  
<http://proceedings.mlr.press/v97/memoli19a.html>

**AUTHORS:** Facundo Memoli, Zane Smith, Zhengchao Wan

**HIGHLIGHT:** We introduce the Wasserstein transform, a method for enhancing and denoising datasets defined on general metric spaces.

453, **TITLE:** Ithemal: Accurate, Portable and Fast Basic Block Throughput Estimation using Deep Neural Networks  
<http://proceedings.mlr.press/v97/mendis19a.html>

**AUTHORS:** Charith Mendis, Alex Renda, Dr.Saman Amarasinghe, Michael Carbin

**HIGHLIGHT:** In this paper we present Ithemal, the first tool which learns to predict the throughput of a set of instructions.

454, **TITLE:** Geometric Losses for Distributional Learning  
<http://proceedings.mlr.press/v97/mensch19a.html>

**AUTHORS:** Arthur Mensch, Mathieu Blondel, Gabriel Peyr?

**HIGHLIGHT:** Building upon recent advances in entropy-regularized optimal transport, and upon Fenchel duality between measures and continuous functions, we propose a generalization of the logistic loss that incorporates a metric or cost between classes.

455, TITLE: Spectral Clustering of Signed Graphs via Matrix Power Means  
<http://proceedings.mlr.press/v97/mercado19a.html>

AUTHORS: Pedro Mercado, Francesco Tudisco, Matthias Hein

**HIGHLIGHT:** We provide a thorough analysis of the proposed approach in the setting of a general Stochastic Block Model that includes models such as the Labeled Stochastic Block Model and the Censored Block Model.

456, TITLE: Simple Stochastic Gradient Methods for Non-Smooth Non-Convex Regularized Optimization  
<http://proceedings.mlr.press/v97/metel19a.html>

AUTHORS: Michael Metel, Akiko Takeda

**HIGHLIGHT:** We present two simple stochastic gradient algorithms, for finite-sum and general stochastic optimization problems, which have superior convergence complexities compared to the current state-of-the-art.

457, TITLE: Reinforcement Learning in Configurable Continuous Environments  
<http://proceedings.mlr.press/v97/metelli19a.html>

AUTHORS: Alberto Maria Metelli, Emanuele Ghelfi, Marcello Restelli

**HIGHLIGHT:** In this paper, we fill this gap by proposing a trust-region method, Relative Entropy Model Policy Search (REMPS), able to learn both the policy and the MDP configuration in continuous domains without requiring the knowledge of the true model of the environment.

458, TITLE: Understanding and correcting pathologies in the training of learned optimizers  
<http://proceedings.mlr.press/v97/metz19a.html>

AUTHORS: Luke Metz, Niru Maheswaranathan, Jeremy Nixon, Daniel Freeman, Jascha Sohl-Dickstein

**HIGHLIGHT:** In this work we propose a training scheme which overcomes both of these difficulties, by dynamically weighting two unbiased gradient estimators for a variational loss on optimizer performance.

459, TITLE: Optimality Implies Kernel Sum Classifiers are Statistically Efficient  
<http://proceedings.mlr.press/v97/meyer19a.html>

AUTHORS: Raphael Meyer, Jean Honorio

**HIGHLIGHT:** We propose a novel combination of optimization tools with learning theory bounds in order to analyze the sample complexity of optimal kernel sum classifiers.

460, TITLE: On Dropout and Nuclear Norm Regularization  
<http://proceedings.mlr.press/v97/mianjy19a.html>

AUTHORS: Poorya Mianjy, Raman Arora

**HIGHLIGHT:** We give a formal and complete characterization of the explicit regularizer induced by dropout in deep linear networks with squared loss.

461, TITLE: Discriminative Regularization for Latent Variable Models with Applications to Electrocardiography  
<http://proceedings.mlr.press/v97/miller19a.html>

AUTHORS: Andrew Miller, Ziad Obermeyer, John Cunningham, Sendhil Mullainathan

AUTHORS: Andrew Miller, Ziad Obermeyer, John Cunningham, Sendhil Mullainathan

**HIGHLIGHT:** We propose a generative model training objective that uses a black-box discriminative model as a regularizer to learn representations that preserve this predictive variation.

462, TITLE: Formal Privacy for Functional Data with Gaussian Perturbations  
<http://proceedings.mlr.press/v97/mirshani19a.html>

AUTHORS: Ardalan Mirshani, Matthew Reimherr, Aleksandra Slavkovic

**HIGHLIGHT:** Motivated by the rapid rise in statistical tools in Functional Data Analysis, we consider the Gaussian mechanism for achieving differential privacy (DP) with parameter estimates taking values in a, potentially infinite-dimensional, separable Banach space.

- 463, TITLE: Co-manifold learning with missing data  
<http://proceedings.mlr.press/v97/mishne19a.html>  
AUTHORS: Gal Mishne, Eric Chi, Ronald Coifman  
HIGHLIGHT: We propose utilizing this coupled structure to perform co-manifold learning: uncovering the underlying geometry of both the rows and the columns of a given matrix, where we focus on a missing data setting.
- 464, TITLE: Agnostic Federated Learning  
<http://proceedings.mlr.press/v97/mohri19a.html>  
AUTHORS: Mehryar Mohri, Gary Sivek, Ananda Theertha Suresh  
HIGHLIGHT: Instead, we propose a new framework of agnostic federated learning, where the centralized model is optimized for any target distribution formed by a mixture of the client distributions.
- 465, TITLE: Flat Metric Minimization with Applications in Generative Modeling  
<http://proceedings.mlr.press/v97/mollenhoff19a.html>  
AUTHORS: Thomas M?llenhoff, Daniel Cremers  
HIGHLIGHT: In our theoretical contribution we prove that the flat metric between a parametrized current and a reference current is Lipschitz continuous in the parameters.
- 466, TITLE: Parsimonious Black-Box Adversarial Attacks via Efficient Combinatorial Optimization  
<http://proceedings.mlr.press/v97/moon19a.html>  
AUTHORS: Seungyong Moon, Gaon An, Hyun Oh Song  
HIGHLIGHT: We propose an efficient discrete surrogate to the optimization problem which does not require estimating the gradient and consequently becomes free of the first order update hyperparameters to tune.
- 467, TITLE: Parameter efficient training of deep convolutional neural networks by dynamic sparse reparameterization  
<http://proceedings.mlr.press/v97/mostafa19a.html>  
AUTHORS: Hesham Mostafa, Xin Wang  
HIGHLIGHT: Here we present a novel dynamic sparse reparameterization method that addresses the limitations of previous techniques such as high computational cost and the need for manual configuration of the number of free parameters allocated to each layer.
- 468, TITLE: A Dynamical Systems Perspective on Nesterov Acceleration  
<http://proceedings.mlr.press/v97/muehlebach19a.html>  
AUTHORS: Michael Muehlebach, Michael Jordan  
HIGHLIGHT: We present a dynamical system framework for understanding Nesterov’s accelerated gradient method.
- 469, TITLE: Relational Pooling for Graph Representations  
<http://proceedings.mlr.press/v97/murphy19a.html>  
AUTHORS: Ryan Murphy, Balasubramaniam Srinivasan, Vinayak Rao, Bruno Ribeiro  
HIGHLIGHT: This work generalizes graph neural networks (GNNs) beyond those based on the Weisfeiler-Lehman (WL) algorithm, graph Laplacians, and diffusions.
- 470, TITLE: Learning Optimal Fair Policies  
<http://proceedings.mlr.press/v97/nabi19a.html>  
AUTHORS: Razieh Nabi, Daniel Malinsky, Ilya Shpitser  
HIGHLIGHT: In this paper, we consider how to make optimal but fair decisions, which “break the cycle of injustice” by correcting for the unfair dependence of both decisions and outcomes on sensitive features (e.g., variables that correspond to gender, race, disability, or other protected attributes).



- 471, TITLE: Lexicographic and Depth-Sensitive Margins in Homogeneous and Non-Homogeneous Deep Models  
<http://proceedings.mlr.press/v97/nacson19a.html>  
AUTHORS: Mor Shpigel Nacson, Suriya Gunasekar, Jason Lee, Nathan Srebro, Daniel Soudry  
HIGHLIGHT: For non-homogeneous ensemble models, which output is a sum of homogeneous sub-models, we show that this solution discards the shallowest sub-models if they are unnecessary.
- 472, TITLE: A Wrapped Normal Distribution on Hyperbolic Space for Gradient-Based Learning  
<http://proceedings.mlr.press/v97/nagano19a.html>  
AUTHORS: Yoshihiro Nagano, Shoichiro Yamaguchi, Yasuhiro Fujita, Masanori Koyama  
HIGHLIGHT: In this paper, we present a novel hyperbolic distribution called hyperbolic wrapped distribution, a wrapped normal distribution on hyperbolic space whose density can be evaluated analytically and differentiated with respect to the parameters.
- 473, TITLE: SGD without Replacement: Sharper Rates for General Smooth Convex Functions  
<http://proceedings.mlr.press/v97/nagaraj19a.html>  
AUTHORS: Dheeraj Nagaraj, Prateek Jain, Praneeth Netrapalli  
HIGHLIGHT: We study stochastic gradient descent without replacement (SGDo) for smooth convex functions.
- 474, TITLE: Dropout as a Structured Shrinkage Prior  
<http://proceedings.mlr.press/v97/nalisnick19a.html>  
AUTHORS: Eric Nalisnick, Jose Miguel Hernandez-Lobato, Padhraic Smyth  
HIGHLIGHT: We propose a novel framework for understanding multiplicative noise in neural networks, considering continuous distributions as well as Bernoulli noise (i.e. dropout).
- 475, TITLE: Hybrid Models with Deep and Invertible Features  
<http://proceedings.mlr.press/v97/nalisnick19b.html>  
AUTHORS: Eric Nalisnick, Akihiro Matsukawa, Yee Whye Teh, Dilan Gorur, Balaji Lakshminarayanan  
HIGHLIGHT: We propose a neural hybrid model consisting of a linear model defined on a set of features computed by a deep, invertible transformation (i.e. a normalizing flow).
- 476, TITLE: Learning Context-dependent Label Permutations for Multi-label Classification  
<http://proceedings.mlr.press/v97/nam19a.html>  
AUTHORS: Jinseok Nam, Young-Bum Kim, Eneldo Loza Mencía, Sunghyun Park, Ruhi Sarikaya, Johannes F?mkranz  
HIGHLIGHT: In this work, we propose a multi-label classification approach which allows to choose a dynamic, context-dependent label ordering.
- 477, TITLE: Zero-Shot Knowledge Distillation in Deep Networks  
<http://proceedings.mlr.press/v97/nayak19a.html>  
AUTHORS: Gaurav Kumar Nayak, Konda Reddy Mopuri, Vaisakh Shaj, Venkatesh Babu Radhakrishnan, Anirban Chakraborty  
HIGHLIGHT: Hence, in this paper, we propose a novel data-free method to train the Student from the Teacher.
- 478, TITLE: A Framework for Bayesian Optimization in Embedded Subspaces  
<http://proceedings.mlr.press/v97/nayebi19a.html>  
AUTHORS: Amin Nayebi, Alexander Munteanu, Matthias Poloczek  
HIGHLIGHT: We present a theoretically founded approach for high-dimensional Bayesian optimization based on low-dimensional subspace embeddings.
- 479, TITLE: Phaseless PCA: Low-Rank Matrix Recovery from Column-wise Phaseless Measurements  
<http://proceedings.mlr.press/v97/nayer19a.html>  
AUTHORS: Seyedehsara Nayer, Praneeth Narayanamurthy, Namrata Vaswani

**HIGHLIGHT:** We introduce a simple algorithm that is provably correct as long as the subspace changes are piecewise constant.  
This work proposes the first set of simple, practically useful, and provable algorithms for two inter-related problems.

480, TITLE: Safe Grid Search with Optimal Complexity  
<http://proceedings.mlr.press/v97/ndiaye19a.html>

AUTHORS: Eugene Ndiaye, Tam Le, Olivier Fercoq, Joseph Salmon, Ichiro Takeuchi  
**HIGHLIGHT:** In this paper, we revisit the techniques of approximating the regularization path up to predefined tolerance  $\epsilon$  in a unified framework and show that its complexity is  $\mathcal{O}(1/\sqrt{d\epsilon})$  for uniformly convex loss of order  $d \geq 2$  and  $\mathcal{O}(1/\sqrt{\epsilon})$  for Generalized Self-Concordant functions.

481, TITLE: Learning to bid in revenue-maximizing auctions  
<http://proceedings.mlr.press/v97/nedelec19a.html>

AUTHORS: Thomas Nedelec, Noureddine El Karoui, Vianney Perchet  
**HIGHLIGHT:** Using a variational approach, we study the complexity of the original objective and we introduce a relaxation of the objective functional in order to use gradient descent methods.

482, TITLE: On Connected Sublevel Sets in Deep Learning  
<http://proceedings.mlr.press/v97/nguyen19a.html>

AUTHORS: Quynh Nguyen  
**HIGHLIGHT:** This paper shows that every sublevel set of the loss function of a class of deep over-parameterized neural nets with piecewise linear activation functions is connected and unbounded.

483, TITLE: Anomaly Detection With Multiple-Hypotheses Predictions  
<http://proceedings.mlr.press/v97/nguyen19b.html>

AUTHORS: Duc Tam Nguyen, Zhongyu Lou, Michael Klar, Thomas Brox  
**HIGHLIGHT:** We propose to learn the data distribution of the foreground more efficiently with a multi-hypotheses autoencoder.

484, TITLE: Non-Asymptotic Analysis of Fractional Langevin Monte Carlo for Non-Convex Optimization  
<http://proceedings.mlr.press/v97/nguyen19c.html>

AUTHORS: Than Huy Nguyen, Umut Simsekli, Gael Richard  
**HIGHLIGHT:** In this study, we analyze the non-asymptotic behavior of FLMC for non-convex optimization and prove finite-time bounds for its expected suboptimality.

485, TITLE: Rotation Invariant Householder Parameterization for Bayesian PCA  
<http://proceedings.mlr.press/v97/nirwan19a.html>

AUTHORS: Rajbir Nirwan, Nils Bertschinger  
**HIGHLIGHT:** Here, we propose a parameterization based on Householder transformations, which remove the rotational symmetry of the posterior.

486, TITLE: Lossless or Quantized Boosting with Integer Arithmetic  
<http://proceedings.mlr.press/v97/nock19a.html>

AUTHORS: Richard Nock, Robert Williamson  
**HIGHLIGHT:** We build a learning algorithm which is able, under mild assumptions, to achieve a lossless boosting-compliant training.

487, TITLE: Training Neural Networks with Local Error Signals  
<http://proceedings.mlr.press/v97/nokland19a.html>

AUTHORS: Arild Nøklund, Lars Hiller Eidnes  
**HIGHLIGHT:** In this paper we demonstrate, for the first time, that layer-wise training can approach the state-of-the-art on a variety of image datasets.

- 488, TITLE: Remember and Forget for Experience Replay  
<http://proceedings.mlr.press/v97/novati19a.html>  
AUTHORS: Guido Novati, Petros Koumoutsakos  
HIGHLIGHT: We introduce Remember and Forget Experience Replay (ReF-ER), a novel method that can enhance RL algorithms with parameterized policies.
- 489, TITLE: Learning to Infer Program Sketches  
<http://proceedings.mlr.press/v97/nye19a.html>  
AUTHORS: Maxwell Nye, Luke Hewitt, Joshua Tenenbaum, Armando Solar-Lezama  
HIGHLIGHT: The key idea of this work is that a flexible combination of pattern recognition and explicit reasoning can be used to solve these complex programming problems.
- 490, TITLE: Tensor Variable Elimination for Plated Factor Graphs  
<http://proceedings.mlr.press/v97/obermeyer19a.html>  
AUTHORS: Fritz Obermeyer, Eli Bingham, Martin Jankowiak, Neeraj Pradhan, Justin Chiu, Alexander Rush, Noah Goodman  
HIGHLIGHT: To exploit efficient tensor algebra in graphs with plates of variables, we generalize undirected factor graphs to plated factor graphs and variable elimination to a tensor variable elimination algorithm that operates directly on plated factor graphs.
- 491, TITLE: Counterfactual Off-Policy Evaluation with Gumbel-Max Structural Causal Models  
<http://proceedings.mlr.press/v97/oberst19a.html>  
AUTHORS: Michael Oberst, David Sontag  
HIGHLIGHT: In particular, we introduce a class of structural causal models (SCMs) for generating counterfactual trajectories in finite partially observable Markov Decision Processes (POMDPs).
- 492, TITLE: Model Function Based Conditional Gradient Method with Armijo-like Line Search  
<http://proceedings.mlr.press/v97/ochs19a.html>  
AUTHORS: Peter Ochs, Yura Malitsky  
HIGHLIGHT: As special cases, for example, we develop an algorithm for additive composite problems and an algorithm for non-linear composite problems which leads to a Gauss-Newton-type algorithm.
- 493, TITLE: TensorFuzz: Debugging Neural Networks with Coverage-Guided Fuzzing  
<http://proceedings.mlr.press/v97/odena19a.html>  
AUTHORS: Augustus Odena, Catherine Olsson, David Andersen, Ian Goodfellow  
HIGHLIGHT: We introduce testing techniques for neural networks that can discover errors occurring only for rare inputs.
- 494, TITLE: Scalable Learning in Reproducing Kernel Krein Spaces  
<http://proceedings.mlr.press/v97/oglic19a.html>  
AUTHORS: Dino Oglic, Thomas G?rtner  
HIGHLIGHT: We provide the first mathematically complete derivation of the Nystr{ö}m method for low-rank approximation of indefinite kernels and propose an efficient method for finding an approximate eigendecomposition of such kernel matrices.
- 495, TITLE: Approximation and non-parametric estimation of ResNet-type convolutional neural networks  
<http://proceedings.mlr.press/v97/oono19a.html>  
AUTHORS: Kenta Oono, Taiji Suzuki  
HIGHLIGHT: We show a ResNet-type CNN can attain the minimax optimal error rates in these classes in more plausible situations – it can be dense, and its width, channel size, and filter size are constant with respect to sample size.
- 496, TITLE: Orthogonal Random Forest for Causal Inference  
<http://proceedings.mlr.press/v97/oprescu19a.html>  
AUTHORS: Miruna Oprescu, Vasilis Syrgkanis, Zhiwei Steven Wu

**HIGHLIGHT:** We propose the orthogonal random forest, an algorithm that combines Neyman-orthogonality to reduce sensitivity with respect to estimation error of nuisance parameters with generalized random forests (Athey et al., 2017)—a flexible non-parametric method for statistical estimation of conditional moment models using random forests.

497, **TITLE:** Inferring Heterogeneous Causal Effects in Presence of Spatial Confounding  
<http://proceedings.mlr.press/v97/osama19a.html>

**AUTHORS:** Muhammad Osama, Dave Zachariah, Thomas Sch?n

**HIGHLIGHT:** We address the problem of inferring the causal effect of an exposure on an outcome across space, using observational data.

498, **TITLE:** Overparameterized Nonlinear Learning: Gradient Descent Takes the Shortest Path?  
<http://proceedings.mlr.press/v97/oymak19a.html>

**AUTHORS:** Samet Oymak, Mahdi Soltanolkotabi

**HIGHLIGHT:** In this paper we demonstrate that when the loss has certain properties over a minimally small neighborhood of the initial point, first order methods such as (stochastic) gradient descent have a few intriguing properties: (1) the iterates converge at a geometric rate to a global optima even when the loss is nonconvex, (2) among all global optima of the loss the iterates converge to one with a near minimal distance to the initial point, (3) the iterates take a near direct route from the initial point to this global optimum.

499, **TITLE:** Multiplicative Weights Updates as a distributed constrained optimization algorithm: Convergence to second-order stationary points almost always

<http://proceedings.mlr.press/v97/panageas19a.html>

**AUTHORS:** Ioannis Panageas, Georgios Piliouras, Xiao Wang

**HIGHLIGHT:** In this paper we focus on constrained non-concave maximization.

500, **TITLE:** Improving Adversarial Robustness via Promoting Ensemble Diversity

<http://proceedings.mlr.press/v97/pang19a.html>

**AUTHORS:** Tianyu Pang, Kun Xu, Chao Du, Ning Chen, Jun Zhu

**HIGHLIGHT:** This paper presents a new method that explores the interaction among individual networks to improve robustness for ensemble models.

501, **TITLE:** Nonparametric Bayesian Deep Networks with Local Competition

<http://proceedings.mlr.press/v97/panousis19a.html>

**AUTHORS:** Konstantinos Panousis, Sotirios Chatzis, Sergios Theodoridis

**HIGHLIGHT:** The aim of this work is to enable inference of deep networks that retain high accuracy for the least possible model complexity, with the latter deduced from the data during inference.

502, **TITLE:** Optimistic Policy Optimization via Multiple Importance Sampling

<http://proceedings.mlr.press/v97/papini19a.html>

**AUTHORS:** Matteo Papini, Alberto Maria Metelli, Lorenzo Lupo, Marcello Restelli

**HIGHLIGHT:** In this paper, we address the exploration-exploitation trade-off in PS by proposing an approach based on Optimism in the Face of Uncertainty.

503, **TITLE:** Deep Residual Output Layers for Neural Language Generation

<http://proceedings.mlr.press/v97/pappas19a.html>

**AUTHORS:** Nikolaos Pappas, James Henderson

**HIGHLIGHT:** In this paper, we investigate the usefulness of more powerful shared mappings for output labels, and propose a deep residual output mapping with dropout between layers to better capture the structure of the output space and avoid overfitting.

504, **TITLE:** Measurements of Three-Level Hierarchical Structure in the Outliers in the Spectrum of Deepnet Hessians

<http://proceedings.mlr.press/v97/papayan19a.html>

**AUTHORS:** Vardan Papyan

**HIGHLIGHT:** We show this term is not a Covariance but a second moment matrix, i.e., it is influenced by means of gradients.

505, **TITLE:** Generalized Majorization-Minimization  
<http://proceedings.mlr.press/v97/parizi19a.html>  
**AUTHORS:** Sobhan Naderi Parizi, Kun He, Reza Aghajani, Stan Sclaroff, Pedro Felzenszwalb  
**HIGHLIGHT:** We generalize MM by relaxing this constraint, and propose a new optimization framework, named Generalized Majorization-Minimization (G-MM), that is more flexible.

506, **TITLE:** Variational Laplace Autoencoders  
<http://proceedings.mlr.press/v97/park19a.html>  
**AUTHORS:** Yookoon Park, Chris Kim, Gunhee Kim  
**HIGHLIGHT:** We present a novel approach that addresses both challenges.

507, **TITLE:** The Effect of Network Width on Stochastic Gradient Descent and Generalization: an Empirical Study  
<http://proceedings.mlr.press/v97/park19b.html>  
**AUTHORS:** Daniel Park, Jascha Sohl-Dickstein, Quoc Le, Samuel Smith  
**HIGHLIGHT:** We investigate how the final parameters found by stochastic gradient descent are influenced by over-parameterization.

508, **TITLE:** Spectral Approximate Inference  
<http://proceedings.mlr.press/v97/park19c.html>  
**AUTHORS:** Sejun Park, Eunho Yang, Se-Young Yun, Jinwoo Shin  
**HIGHLIGHT:** To overcome the limitation, we propose a novel approach utilizing the global spectral feature of GM.

509, **TITLE:** Self-Supervised Exploration via Disagreement  
<http://proceedings.mlr.press/v97/pathak19a.html>  
**AUTHORS:** Deepak Pathak, Dhiraj Gandhi, Abhinav Gupta  
**HIGHLIGHT:** In this paper, we propose a formulation for exploration inspired by the work in active learning literature.

510, **TITLE:** Subspace Robust Wasserstein Distances  
<http://proceedings.mlr.press/v97/paty19a.html>  
**AUTHORS:** Fran?ois-Pierre Paty, Marco Cuturi  
**HIGHLIGHT:** We propose in this work a “max-min” robust variant of the Wasserstein distance by considering the maximal possible distance that can be realized between two measures, assuming they can be projected orthogonally on a lower k-dimensional subspace.

511, **TITLE:** Fingerprint Policy Optimisation for Robust Reinforcement Learning  
<http://proceedings.mlr.press/v97/paul19a.html>  
**AUTHORS:** Supratik Paul, Michael A. Osborne, Shimon Whiteson  
**HIGHLIGHT:** In this paper, we present fingerprint policy optimisation (FPO), which finds a policy that is optimal in expectation across the distribution of environment variables.

512, **TITLE:** COMIC: Multi-view Clustering Without Parameter Selection  
<http://proceedings.mlr.press/v97/peng19a.html>  
**AUTHORS:** Xi Peng, Zhenyu Huang, Jiancheng Lv, Hongyuan Zhu, Joey Tianyi Zhou  
**HIGHLIGHT:** In this paper, we study two challenges in clustering analysis, namely, how to cluster multi-view data and how to perform clustering without parameter selection on cluster size.

513, **TITLE:** Domain Agnostic Learning with Disentangled Representations

<http://proceedings.mlr.press/v97/peng19b.html>

AUTHORS: Xingchao Peng, Zijun Huang, Ximeng Sun, Kate Saenko  
HIGHLIGHT: In this paper, we propose the task of Domain-Agnostic Learning (DAL): How to transfer knowledge from a labeled source domain to unlabeled data from arbitrary target domains?

514, TITLE: Collaborative Channel Pruning for Deep Networks

<http://proceedings.mlr.press/v97/peng19c.html>

AUTHORS: Hanyu Peng, Jiayang Wu, Shifeng Chen, Junzhou Huang  
HIGHLIGHT: In this paper, we propose a novel algorithm, namely collaborative channel pruning (CCP), to reduce the computational overhead with negligible performance degradation.

515, TITLE: Exploiting structure of uncertainty for efficient matroid semi-bandits

<http://proceedings.mlr.press/v97/perrault19a.html>

AUTHORS: Pierre Perrault, Vianney Perchet, Michal Valko  
HIGHLIGHT: We improve the efficiency of algorithms for stochastic combinatorial semi-bandits.

516, TITLE: Cognitive model priors for predicting human decisions

<http://proceedings.mlr.press/v97/peterson19a.html>

AUTHORS: Joshua Peterson, David Bourgin, Daniel Reichman, Thomas Griffiths, Stuart Russell  
HIGHLIGHT: We argue that this is mainly due to data scarcity, since noisy human behavior requires massive sample sizes to be accurately captured by off-the-shelf machine learning methods.  
Second, we present the first large-scale dataset for human decision-making, containing over 240,000 human judgments across over 13,000 decision problems.

517, TITLE: Towards Understanding Knowledge Distillation

<http://proceedings.mlr.press/v97/phuong19a.html>

AUTHORS: Mary Phuong, Christoph Lampert  
HIGHLIGHT: In this work, we provide the first insights into the working mechanisms of distillation by studying the special case of linear and deep linear classifiers.

518, TITLE: Temporal Gaussian Mixture Layer for Videos

<http://proceedings.mlr.press/v97/piergiovanni19a.html>

AUTHORS: Aj Piergiovanni, Michael Ryoo  
HIGHLIGHT: We present our fully convolutional video models with multiple TGM layers for activity detection.

519, TITLE: Voronoi Boundary Classification: A High-Dimensional Geometric Approach via Weighted Monte Carlo Integration

<http://proceedings.mlr.press/v97/polianskii19a.html>

AUTHORS: Vladislav Polianskii, Florian T. Pokorný  
HIGHLIGHT: We propose a Monte-Carlo integration based approach that instead computes a weighted integral over the boundaries of Voronoi cells, thus incorporating additional information about the Voronoi cell structure.

520, TITLE: On Variational Bounds of Mutual Information

<http://proceedings.mlr.press/v97/poole19a.html>

AUTHORS: Ben Poole, Sherjil Ozair, Aaron Van Den Oord, Alex Alemi, George Tucker  
HIGHLIGHT: In this work, we unify these recent developments in a single framework.

521, TITLE: Hiring Under Uncertainty

<http://proceedings.mlr.press/v97/purohit19a.html>

AUTHORS: Manish Purohit, Sreenivas Gollapudi, Manish Raghavan  
HIGHLIGHT: In this paper we introduce the hiring under uncertainty problem to model the questions faced by hiring committees in large enterprises and universities alike.

- 522, TITLE: SAGA with Arbitrary Sampling  
<http://proceedings.mlr.press/v97/qian19a.html>  
AUTHORS: Xun Qian, Zheng Qu, Peter Richtarik  
HIGHLIGHT: We remedy this situation and propose a general and flexible variant of SAGA following the arbitrary sampling paradigm.
- 523, TITLE: SGD with Arbitrary Sampling: General Analysis and Improved Rates  
<http://proceedings.mlr.press/v97/qian19b.html>  
AUTHORS: Xun Qian, Peter Richtarik, Robert Gower, Alibek Sailanbayev, Nicolas Loizou, Egor Shulgin  
HIGHLIGHT: We propose a general yet simple theorem describing the convergence of SGD under the arbitrary sampling paradigm.
- 524, TITLE: AutoVC: Zero-Shot Voice Style Transfer with Only Autoencoder Loss  
<http://proceedings.mlr.press/v97/qian19c.html>  
AUTHORS: Kaizhi Qian, Yang Zhang, Shiyu Chang, Xuesong Yang, Mark Hasegawa-Johnson  
HIGHLIGHT: In this paper, we propose a new style transfer scheme that involves only an autoencoder with a carefully designed bottleneck.
- 525, TITLE: Fault Tolerance in Iterative-Convergent Machine Learning  
<http://proceedings.mlr.press/v97/qiao19a.html>  
AUTHORS: Aurick Qiao, Bryon Aragam, Bingjing Zhang, Eric Xing  
HIGHLIGHT: In this paper, we develop a general framework to quantify the effects of calculation errors on iterative-convergent algorithms.
- 526, TITLE: Imperceptible, Robust, and Targeted Adversarial Examples for Automatic Speech Recognition  
<http://proceedings.mlr.press/v97/qin19a.html>  
AUTHORS: Yao Qin, Nicholas Carlini, Garrison Cottrell, Ian Goodfellow, Colin Raffel  
HIGHLIGHT: This paper makes progress on both of these fronts.
- 527, TITLE: GMNN: Graph Markov Neural Networks  
<http://proceedings.mlr.press/v97/qu19a.html>  
AUTHORS: Meng Qu, Yoshua Bengio, Jian Tang  
HIGHLIGHT: In this paper, we propose the Graph Markov Neural Network (GMNN) that combines the advantages of both worlds.
- 528, TITLE: Nonlinear Distributional Gradient Temporal-Difference Learning  
<http://proceedings.mlr.press/v97/qu19b.html>  
AUTHORS: Chao Qu, Shie Mannor, Huan Xu  
HIGHLIGHT: In the control setting, we propose the distributional Greedy-GQ using similar derivation.
- 529, TITLE: Learning to Collaborate in Markov Decision Processes  
<http://proceedings.mlr.press/v97/radanovic19a.html>  
AUTHORS: Goran Radanovic, Rati Devidze, David Parkes, Adish Singla  
HIGHLIGHT: We consider a two-agent MDP framework where agents repeatedly solve a task in a collaborative setting.
- 530, TITLE: Meta-Learning Neural Bloom Filters  
<http://proceedings.mlr.press/v97/rae19a.html>  
AUTHORS: Jack Rae, Sergey Bartunov, Timothy Lillicrap  
HIGHLIGHT: In this paper we explore the learning of approximate set membership over a set of data in one-shot via meta-learning.
- 531, TITLE: Direct Uncertainty Prediction for Medical Second Opinions

<http://proceedings.mlr.press/v97/raghu19a.html>

AUTHORS: Maithra Raghu, Katy Blumer, Rory Sayres, Ziad Obermeyer, Bobby Kleinberg, Sendhil Mullainathan, Jon Kleinberg

HIGHLIGHT: In this work, we show that machine learning models can be successfully trained to give uncertainty scores to data instances that result in high expert disagreements.

532, TITLE: Game Theoretic Optimization via Gradient-based Nikaido-Isoda Function

<http://proceedings.mlr.press/v97/raghunathan19a.html>

AUTHORS: Arvind Raghunathan, Anoop Cherian, Devesh Jha

HIGHLIGHT: To this end, we introduce the Gradient-based Nikaido-Isoda (GNI) function which serves: (i) as a merit function, vanishing only at the first-order stationary points of each player's optimization problem, and (ii) provides error bounds to a stationary Nash point.

533, TITLE: On the Spectral Bias of Neural Networks

<http://proceedings.mlr.press/v97/rahaman19a.html>

AUTHORS: Nasim Rahaman, Aristide Baratin, Devansh Arpit, Felix Draxler, Min Lin, Fred Hamprecht, Yoshua Bengio, Aaron Courville

HIGHLIGHT: In this work we present properties of neural networks that complement this aspect of expressivity.

534, TITLE: Look Ma, No Latent Variables: Accurate Cutset Networks via Compilation

<http://proceedings.mlr.press/v97/rahman19a.html>

AUTHORS: Tahrima Rahman, Shasha Jin, Vibhav Gogate

HIGHLIGHT: To address this problem, in this paper, we propose a novel approach for inducing cutset networks, a well-known tractable, highly interpretable representation that does not use latent variables and admits linear time MAR as well as MAP inference.

535, TITLE: Does Data Augmentation Lead to Positive Margin?

<http://proceedings.mlr.press/v97/rajput19a.html>

AUTHORS: Shashank Rajput, Zhili Feng, Zachary Charles, Po-Ling Loh, Dimitris Papailiopoulos

HIGHLIGHT: In this work, we analyze the robustness that DA begets by quantifying the margin that DA enforces on empirical risk minimizers.

536, TITLE: Efficient Off-Policy Meta-Reinforcement Learning via Probabilistic Context Variables

<http://proceedings.mlr.press/v97/rakelly19a.html>

AUTHORS: Kate Rakelly, Aurick Zhou, Chelsea Finn, Sergey Levine, Deirdre Quillen

HIGHLIGHT: In this paper, we address these challenges by developing an off-policy meta-RL algorithm that disentangles task inference and control.

537, TITLE: Screening rules for Lasso with non-convex Sparse Regularizers

<http://proceedings.mlr.press/v97/rakotomamonjy19a.html>

AUTHORS: Alain Rakotomamonjy, Gilles Gasso, Joseph Salmon

HIGHLIGHT: The approach we propose is based on an iterative majorization-minimization (MM) strategy that includes a screening rule in the inner solver and a condition for propagating screened variables between iterations of MM.

538, TITLE: Topological Data Analysis of Decision Boundaries with Application to Model Selection

<http://proceedings.mlr.press/v97/ramamurthy19a.html>

AUTHORS: Karthikeyan Natesan Ramamurthy, Kush Varshney, Krishnan Mody

HIGHLIGHT: We propose the labeled Čech complex, the plain labeled Vietoris-Rips complex, and the locally scaled labeled Vietoris-Rips complex to perform persistent homology inference of decision boundaries in classification tasks.

539, TITLE: HyperGAN: A Generative Model for Diverse, Performant Neural Networks

<http://proceedings.mlr.press/v97/ratzlaff19a.html>



- AUTHORS: Neale Ratzlaff, Li Fuxin  
HIGHLIGHT: We introduce HyperGAN, a generative model that learns to generate all the parameters of a deep neural network.
- 540, TITLE: Efficient On-Device Models using Neural Projections  
<http://proceedings.mlr.press/v97/ravi19a.html>  
AUTHORS: Sujith Ravi  
HIGHLIGHT: We propose a neural projection approach for training compact on-device neural networks.
- 541, TITLE: A Block Coordinate Descent Proximal Method for Simultaneous Filtering and Parameter Estimation  
<http://proceedings.mlr.press/v97/razi-perchikolaei19a.html>  
AUTHORS: Ramin Razi-perchikolaei, Harish Bhat  
HIGHLIGHT: We propose and analyze a block coordinate descent proximal algorithm (BCD-prox) for simultaneous filtering and parameter estimation of ODE models.
- 542, TITLE: Do ImageNet Classifiers Generalize to ImageNet?  
<http://proceedings.mlr.press/v97/recht19a.html>  
AUTHORS: Benjamin Recht, Rebecca Roelofs, Ludwig Schmidt, Vaishal Shankar  
HIGHLIGHT: We build new test sets for the CIFAR-10 and ImageNet datasets.
- 543, TITLE: Fast Rates for a kNN Classifier Robust to Unknown Asymmetric Label Noise  
<http://proceedings.mlr.press/v97/reeve19a.html>  
AUTHORS: Henry Reeve, Ata Kaban  
HIGHLIGHT: We consider classification in the presence of class-dependent asymmetric label noise with unknown noise probabilities.
- 544, TITLE: Almost Unsupervised Text to Speech and Automatic Speech Recognition  
<http://proceedings.mlr.press/v97/ren19a.html>  
AUTHORS: Yi Ren, Xu Tan, Tao Qin, Sheng Zhao, Zhou Zhao, Tie-Yan Liu  
HIGHLIGHT: In this paper, by leveraging the dual nature of the two tasks, we propose an almost unsupervised learning method that only leverages few hundreds of paired data and extra unpaired data for TTS and ASR.
- 545, TITLE: Adaptive Antithetic Sampling for Variance Reduction  
<http://proceedings.mlr.press/v97/ren19b.html>  
AUTHORS: Hongyu Ren, Shengjia Zhao, Stefano Ermon  
HIGHLIGHT: In this paper, we propose a general-purpose adaptive antithetic sampling framework.
- 546, TITLE: Adversarial Online Learning with noise  
<http://proceedings.mlr.press/v97/resler19a.html>  
AUTHORS: Alon Resler, Yishay Mansour  
HIGHLIGHT: Specifically, we consider binary losses XORed with the noise, which is a Bernoulli random variable.
- 547, TITLE: A Polynomial Time MCMC Method for Sampling from Continuous Determinantal Point Processes  
<http://proceedings.mlr.press/v97/rezaei19a.html>  
AUTHORS: Alireza Rezaei, Shayan Oveis Gharan  
HIGHLIGHT: We study the Gibbs sampling algorithm for discrete and continuous  $\mathbb{S}_k$ -determinantal point processes.
- 548, TITLE: A Persistent Weisfeiler-Lehman Procedure for Graph Classification  
<http://proceedings.mlr.press/v97/rieck19a.html>  
AUTHORS: Bastian Rieck, Christian Bock, Karsten Borgwardt

**HIGHLIGHT:** Our method, which we formalise as a generalisation of Weisfeiler–Lehman subtree features, exhibits favourable classification accuracy and its improvements in predictive performance are mainly driven by including cycle information.

549, **TITLE:** Efficient learning of smooth probability functions from Bernoulli tests with guarantees  
<http://proceedings.mlr.press/v97/rolland19a.html>  
**AUTHORS:** Paul Rolland, Ali Kavis, Alexander Immer, Adish Singla, Volkan Cevher  
**HIGHLIGHT:** We study the fundamental problem of learning an unknown, smooth probability function via point-wise Bernoulli tests.

550, **TITLE:** Separable value functions across time-scales  
<http://proceedings.mlr.press/v97/romoff19a.html>  
**AUTHORS:** Joshua Romoff, Peter Henderson, Ahmed Touati, Yann Ollivier, Joelle Pineau, Emma Brunskill  
**HIGHLIGHT:** We present an extension of temporal difference (TD) learning, which we call TD( $\Delta$ ), that breaks down a value function into a series of components based on the differences between value functions with smaller discount factors.

551, **TITLE:** Online Convex Optimization in Adversarial Markov Decision Processes  
<http://proceedings.mlr.press/v97/rosenberg19a.html>  
**AUTHORS:** Aviv Rosenberg, Yishay Mansour  
**HIGHLIGHT:** We consider online learning in episodic loop-free Markov decision processes (MDPs), where the loss function can change arbitrarily between episodes, and the transition function is not known to the learner.

552, **TITLE:** Good Initializations of Variational Bayes for Deep Models  
<http://proceedings.mlr.press/v97/rossi19a.html>  
**AUTHORS:** Simone Rossi, Pietro Michiardi, Maurizio Filippone  
**HIGHLIGHT:** We address this by proposing a novel layer-wise initialization strategy based on Bayesian linear models.

553, **TITLE:** The Odds are Odd: A Statistical Test for Detecting Adversarial Examples  
<http://proceedings.mlr.press/v97/roth19a.html>  
**AUTHORS:** Kevin Roth, Yannic Kilcher, Thomas Hofmann  
**HIGHLIGHT:** We investigate conditions under which test statistics exist that can reliably detect examples, which have been adversarially manipulated in a white-box attack.

554, **TITLE:** Neuron birth-death dynamics accelerates gradient descent and converges asymptotically  
<http://proceedings.mlr.press/v97/rotskoff19a.html>  
**AUTHORS:** Grant Rotskoff, Samy Jelassi, Joan Bruna, Eric Vanden-Eijnden  
**HIGHLIGHT:** In this work, we propose a non-local mass transport dynamics that leads to a modified PDE with the same minimizer.

555, **TITLE:** Iterative Linearized Control: Stable Algorithms and Complexity Guarantees  
<http://proceedings.mlr.press/v97/roulet19a.html>  
**AUTHORS:** Vincent Roulet, Dmitriy Drusvyatskiy, Siddhartha Srinivasa, Zaid Harchaoui  
**HIGHLIGHT:** We examine popular gradient-based algorithms for nonlinear control in the light of the modern complexity analysis of first-order optimization algorithms.

556, **TITLE:** Statistics and Samples in Distributional Reinforcement Learning  
<http://proceedings.mlr.press/v97/rowland19a.html>  
**AUTHORS:** Mark Rowland, Robert Dadashi, Saurabh Kumar, Remi Munos, Marc G. Bellemare, Will Dabney  
**HIGHLIGHT:** We present a unifying framework for designing and analysing distributional reinforcement learning (DRL) algorithms in terms of recursively estimating statistics of the return distribution.

- 557, TITLE: A Contrastive Divergence for Combining Variational Inference and MCMC  
<http://proceedings.mlr.press/v97/ruiz19a.html>  
AUTHORS: Francisco Ruiz, Michalis Titsias  
HIGHLIGHT: To make inference tractable, we introduce the variational contrastive divergence (VCD), a new divergence that replaces the standard Kullback-Leibler (KL) divergence used in VI.
- 558, TITLE: Plug-and-Play Methods Provably Converge with Properly Trained Denoisers  
<http://proceedings.mlr.press/v97/ryu19a.html>  
AUTHORS: Ernest Ryu, Jialin Liu, Sicheng Wang, Xiaohan Chen, Zhangyang Wang, Wotao Yin  
HIGHLIGHT: In this paper, we theoretically establish convergence of PnP-FBS and PnP-ADMM, without using diminishing stepsizes, under a certain Lipschitz condition on the denoisers.
- 559, TITLE: White-box vs Black-box: Bayes Optimal Strategies for Membership Inference  
<http://proceedings.mlr.press/v97/sablayrolles19a.html>  
AUTHORS: Alexandre Sablayrolles, Matthijs Douze, Cordelia Schmid, Yann Ollivier, Herve Jegou  
HIGHLIGHT: In this paper, we derive the optimal strategy for membership inference with a few assumptions on the distribution of the parameters.
- 560, TITLE: Tractable n-Metrics for Multiple Graphs  
<http://proceedings.mlr.press/v97/safavi19a.html>  
AUTHORS: Sam Safavi, Jose Bento  
HIGHLIGHT: In this paper, we introduce a new family of multi-distances (a distance between more than two elements) that satisfies a generalization of the properties of metrics to multiple elements.
- 561, TITLE: An Optimal Private Stochastic-MAB Algorithm based on Optimal Private Stopping Rule  
<http://proceedings.mlr.press/v97/sajed19a.html>  
AUTHORS: Touqir Sajed, Or Sheffét  
HIGHLIGHT: We present a provably optimal differentially private algorithm for the stochastic multi-arm bandit problem, as opposed to the private analogue of the UCB-algorithm (Mishra and Thakurta, 2015; Tossou and Dimitrakakis, 2016) which doesn't meet the recently discovered lower-bound of  $\Omega\left(\frac{K \log(T)}{\epsilon}\right)$  (Shariff and Sheffét, 2018).
- 562, TITLE: Deep Gaussian Processes with Importance-Weighted Variational Inference  
<http://proceedings.mlr.press/v97/salimbeni19a.html>  
AUTHORS: Hugh Salimbeni, Vincent Dutordoir, James Hensman, Marc Deisenroth  
HIGHLIGHT: We instead incorporate noisy variables as latent covariates, and propose a novel importance-weighted objective, which leverages analytic results and provides a mechanism to trade off computation for improved accuracy.
- 563, TITLE: Multivariate Submodular Optimization  
<http://proceedings.mlr.press/v97/santiago19a.html>  
AUTHORS: Richard Santiago, F. Bruce Shepherd  
HIGHLIGHT: In this work we focus on a more general class of multivariate submodular optimization (MVSO) problems:  $\min^{\max} f(S_1, S_2, \dots, S_k) : S_1 \uplus S_2 \uplus \dots \uplus S_k \in \mathcal{F}$ .
- 564, TITLE: Near optimal finite time identification of arbitrary linear dynamical systems  
<http://proceedings.mlr.press/v97/sarkar19a.html>  
AUTHORS: Tuhin Sarkar, Alexander Rakhlin  
HIGHLIGHT: We provide the first analysis of the general case when eigenvalues of the LTI system are arbitrarily distributed in three regimes: stable, marginally stable, and explosive.
- 565, TITLE: Breaking Inter-Layer Co-Adaptation by Classifier Anonymization  
<http://proceedings.mlr.press/v97/sato19a.html>  
AUTHORS: Ikuro Sato, Kohta Ishikawa, Guoqing Liu, Masayuki Tanaka

**HIGHLIGHT:** We introduce a method called Feature-extractor Optimization through Classifier Anonymization (FOCA), which is designed to avoid an explicit co-adaptation between a feature extractor and a particular classifier by using many randomly-generated, weak classifiers during optimization.

566, **TITLE:** A Theoretical Analysis of Contrastive Unsupervised Representation Learning  
<http://proceedings.mlr.press/v97/saunshi19a.html>

**AUTHORS:** Nikunj Saunshi, Orestis Plevrakis, Sanjeev Arora, Mikhail Khodak, Hrishikesh Khandeparkar  
**HIGHLIGHT:** The current paper uses the term contrastive learning for such algorithms and presents a theoretical framework for analyzing them by introducing latent classes and hypothesizing that semantically similar points are sampled from the same latent class.

567, **TITLE:** Locally Private Bayesian Inference for Count Models  
<http://proceedings.mlr.press/v97/schein19a.html>

**AUTHORS:** Aaron Schein, Zhiwei Steven Wu, Alexandra Schofield, Mingyuan Zhou, Hanna Wallach  
**HIGHLIGHT:** We present a general and modular method for privacy-preserving Bayesian inference for Poisson factorization, a broad class of models that includes some of the most widely used models in the social sciences.

568, **TITLE:** Weakly-Supervised Temporal Localization via Occurrence Count Learning  
<http://proceedings.mlr.press/v97/schroeter19a.html>

**AUTHORS:** Julien Schroeter, Kirill Sidorov, David Marshall  
**HIGHLIGHT:** We propose a novel model for temporal detection and localization which allows the training of deep neural networks using only counts of event occurrences as training labels.

569, **TITLE:** Discovering Context Effects from Raw Choice Data  
<http://proceedings.mlr.press/v97/seshadri19a.html>

**AUTHORS:** Arjun Seshadri, Alex Peysakhovich, Johan Ugander  
**HIGHLIGHT:** In this work, our goal is to discover such choice set effects from raw choice data.

570, **TITLE:** On the Feasibility of Learning, Rather than Assuming, Human Biases for Reward Inference  
<http://proceedings.mlr.press/v97/shah19a.html>

**AUTHORS:** Rohin Shah, Noah Gundotra, Pieter Abbeel, Anca Dragan  
**HIGHLIGHT:** Our goal is for agents to optimize the right reward function, despite how difficult it is for us to specify what that is.

571, **TITLE:** Exploration Conscious Reinforcement Learning Revisited  
<http://proceedings.mlr.press/v97/shani19a.html>

**AUTHORS:** Lior Shani, Yonathan Efroni, Shie Mannor  
**HIGHLIGHT:** In this work, we take a different approach and study exploration-conscious criteria, that result in optimal policies with respect to the exploration mechanism.

572, **TITLE:** Compressed Factorization: Fast and Accurate Low-Rank Factorization of Compressively-Sensed Data  
<http://proceedings.mlr.press/v97/sharan19a.html>

**AUTHORS:** Vatsal Sharan, Kai Sheng Tai, Peter Bailis, Gregory Valiant  
**HIGHLIGHT:** In this work, we consider the question of accurately and efficiently computing low-rank matrix or tensor factorizations given data compressed via random projections.

573, **TITLE:** Conditional Independence in Testing Bayesian Networks  
<http://proceedings.mlr.press/v97/shen19a.html>

**AUTHORS:** Yujia Shen, Haiying Huang, Arthur Choi, Adnan Darwiche  
**HIGHLIGHT:** In this paper, we study conditional independence in TBNs, showing that it can be inferred from d-separation as in BNs.

- 574, TITLE: Learning to Clear the Market  
<http://proceedings.mlr.press/v97/shen19b.html>  
AUTHORS: Weiran Shen, Sebastien Lahaie, Renato Paes Leme  
HIGHLIGHT: In this work, we cast the problem of predicting clearing prices into a learning framework and use the resulting models to perform revenue optimization in auctions and markets with contextual information.
- 575, TITLE: Mixture Models for Diverse Machine Translation: Tricks of the Trade  
<http://proceedings.mlr.press/v97/shen19c.html>  
AUTHORS: Tianxiao Shen, Myle Ott, Michael Auli, Marc Aurelio Ranzato  
HIGHLIGHT: Mixture Models for Diverse Machine Translation: Tricks of the Trade.
- 576, TITLE: Hessian Aided Policy Gradient  
<http://proceedings.mlr.press/v97/shen19d.html>  
AUTHORS: Zebang Shen, Alejandro Ribeiro, Hamed Hassani, Hui Qian, Chao Mi  
HIGHLIGHT: This paper presents a Hessian aided policy gradient method with the first improved sample complexity of  $\mathcal{O}M(\frac{1}{\epsilon^3})$ .
- 577, TITLE: Learning with Bad Training Data via Iterative Trimmed Loss Minimization  
<http://proceedings.mlr.press/v97/shen19e.html>  
AUTHORS: Yanyao Shen, Sujay Sanghavi  
HIGHLIGHT: In this paper, we study a simple and generic framework to tackle the problem of learning model parameters when a fraction of the training samples are corrupted.
- 578, TITLE: Replica Conditional Sequential Monte Carlo  
<http://proceedings.mlr.press/v97/shestopaloff19a.html>  
AUTHORS: Alex Shestopaloff, Arnaud Doucet  
HIGHLIGHT: We propose a Markov chain Monte Carlo (MCMC) scheme to perform state inference in non-linear non-Gaussian state-space models.
- 579, TITLE: Scalable Training of Inference Networks for Gaussian-Process Models  
<http://proceedings.mlr.press/v97/shi19a.html>  
AUTHORS: Jiaxin Shi, Mohammad Emtiyaz Khan, Jun Zhu  
HIGHLIGHT: We propose an algorithm that enables such training by tracking a stochastic, functional mirror-descent algorithm.
- 580, TITLE: Fast Direct Search in an Optimally Compressed Continuous Target Space for Efficient Multi-Label Active Learning  
<http://proceedings.mlr.press/v97/shi19b.html>  
AUTHORS: Weishi Shi, Qi Yu  
HIGHLIGHT: We propose a novel CS-BPCA process that integrates compressed sensing and Bayesian principal component analysis to perform a two-level label transformation, resulting in an optimally compressed continuous target space.
- 581, TITLE: Model-Based Active Exploration  
<http://proceedings.mlr.press/v97/shyam19a.html>  
AUTHORS: Pranav Shyam, Wojciech Jaskowski, Faustino Gomez  
HIGHLIGHT: This paper introduces an efficient active exploration algorithm, Model-Based Active eXploration (MAX), which uses an ensemble of forward models to plan to observe novel events.
- 582, TITLE: Rehashing Kernel Evaluation in High Dimensions  
<http://proceedings.mlr.press/v97/siminelakis19a.html>  
AUTHORS: Paris Siminelakis, Kexin Rong, Peter Bailis, Moses Charikar, Philip Levis  
HIGHLIGHT: In this paper, we close the gap between theory and practice by addressing these challenges via provable and practical procedures for adaptive sample size selection, preprocessing time reduction, and refined

variance bounds that quantify the data-dependent performance of random sampling and hashing-based kernel evaluation methods.

- 583, TITLE: Revisiting precision recall definition for generative modeling  
<http://proceedings.mlr.press/v97/simon19a.html>  
AUTHORS: Loic Simon, Ryan Webster, Julien Rabin  
HIGHLIGHT: In this article we revisit the definition of Precision-Recall (PR) curves for generative models proposed by (Sajjadi et al., 2018).
- 584, TITLE: First-Order Adversarial Vulnerability of Neural Networks and Input Dimension  
<http://proceedings.mlr.press/v97/simon-gabriel19a.html>  
AUTHORS: Carl-Johann Simon-Gabriel, Yann Ollivier, Leon Bottou, Bernhard Schölkopf, David Lopez-Paz  
HIGHLIGHT: We show that adversarial vulnerability increases with the gradients of the training objective when viewed as a function of the inputs.
- 585, TITLE: Refined Complexity of PCA with Outliers  
<http://proceedings.mlr.press/v97/simonov19a.html>  
AUTHORS: Kirill Simonov, Fedor Fomin, Petr Golovach, Fahad Panolan  
HIGHLIGHT: We provide a rigorous algorithmic analysis of the problem.
- 586, TITLE: A Tail-Index Analysis of Stochastic Gradient Noise in Deep Neural Networks  
<http://proceedings.mlr.press/v97/simsekli19a.html>  
AUTHORS: Umut Simsekli, Levent Sagun, Mert Gurbuzbalaban  
HIGHLIGHT: Accordingly, we propose to analyze SGD as an SDE driven by a Lévy motion.
- 587, TITLE: Non-Parametric Priors For Generative Adversarial Networks  
<http://proceedings.mlr.press/v97/singh19a.html>  
AUTHORS: Rajhans Singh, Pavan Turaga, Suren Jayasuriya, Ravi Garg, Martin Braun  
HIGHLIGHT: We present a straightforward formalization of this problem; using basic results from probability theory and off-the-shelf-optimization tools, we develop ways to arrive at appropriate non-parametric priors.
- 588, TITLE: Understanding Impacts of High-Order Loss Approximations and Features in Deep Learning Interpretation  
<http://proceedings.mlr.press/v97/singla19a.html>  
AUTHORS: Sahil Singla, Eric Wallace, Shi Feng, Soheil Feizi  
HIGHLIGHT: We use an  $L_0 - L_1$  relaxation technique along with proximal gradient descent to efficiently compute group-feature importance values.
- 589, TITLE: kernelPSI: a Post-Selection Inference Framework for Nonlinear Variable Selection  
<http://proceedings.mlr.press/v97/slim19a.html>  
AUTHORS: Lotfi Slim, Clément Chatelain, Chloe-Agathe Azencott, Jean-Philippe Vert  
HIGHLIGHT: In the present work, we exploit recent advances in post-selection inference to propose a valid statistical test for the association of a joint model of the selected kernels with the outcome.
- 590, TITLE: GEOMETrics: Exploiting Geometric Structure for Graph-Encoded Objects  
<http://proceedings.mlr.press/v97/smith19a.html>  
AUTHORS: Edward Smith, Adriana Romero, Scott Fujimoto, David Meger  
HIGHLIGHT: In this paper, we argue that the graph representation of geometric objects allows for additional structure, which should be leveraged for enhanced reconstruction.
- 591, TITLE: The Evolved Transformer  
<http://proceedings.mlr.press/v97/so19a.html>  
AUTHORS: David So, Quoc Le, Chen Liang

**HIGHLIGHT:** Our goal is to apply NAS to search for a better alternative to the Transformer.

592, **TITLE:** QTRAN: Learning to Factorize with Transformation for Cooperative Multi-Agent Reinforcement Learning

<http://proceedings.mlr.press/v97/son19a.html>

**AUTHORS:** Kyunghwan Son, Daewoo Kim, Wan Ju Kang, David Earl Hostallero, Yung Yi

**HIGHLIGHT:** In this paper, we propose a new factorization method for MARL, QTRAN, which is free from such structural constraints and takes on a new approach to transforming the original joint action-value function into an easily factorizable one, with the same optimal actions.

593, **TITLE:** Distribution calibration for regression

<http://proceedings.mlr.press/v97/song19a.html>

**AUTHORS:** Hao Song, Tom Diethe, Meelis Kull, Peter Flach

**HIGHLIGHT:** We introduce the novel concept of distribution calibration, and demonstrate its advantages over the existing definition of quantile calibration.

594, **TITLE:** SELFIE: Refurbishing Unclean Samples for Robust Deep Learning

<http://proceedings.mlr.press/v97/song19b.html>

**AUTHORS:** Hwanjun Song, Minseok Kim, Jae-Gil Lee

**HIGHLIGHT:** To overcome overfitting on the noisy labels, we propose a novel robust training method called SELFIE.

595, **TITLE:** Revisiting the Softmax Bellman Operator: New Benefits and New Perspective

<http://proceedings.mlr.press/v97/song19c.html>

**AUTHORS:** Zhao Song, Ron Parr, Lawrence Carin

**HIGHLIGHT:** To better understand how and why this occurs, we revisit theoretical properties of the softmax Bellman operator, and prove that (i) it converges to the standard Bellman operator exponentially fast in the inverse temperature parameter, and (ii) the distance of its Q function from the optimal one can be bounded.

596, **TITLE:** MASS: Masked Sequence to Sequence Pre-training for Language Generation

<http://proceedings.mlr.press/v97/song19d.html>

**AUTHORS:** Kaitao Song, Xu Tan, Tao Qin, Jianfeng Lu, Tie-Yan Liu

**HIGHLIGHT:** Inspired by the success of BERT, we propose MAsked Sequence to Sequence pre-training (MASS) for the encoder-decoder based language generation tasks.

597, **TITLE:** Dual Entangled Polynomial Code: Three-Dimensional Coding for Distributed Matrix Multiplication

<http://proceedings.mlr.press/v97/soto19a.html>

**AUTHORS:** Pedro Soto, Jun Li, Xiaodi Fan

**HIGHLIGHT:** In this paper, we propose dual entangled polynomial (DEP) codes that require around 25% fewer tasks than EP codes by executing two matrix multiplications on each task.

598, **TITLE:** Compressing Gradient Optimizers via Count-Sketches

<http://proceedings.mlr.press/v97/spring19a.html>

**AUTHORS:** Ryan Spring, Anastasios Kyrillidis, Vijai Mohan, Anshumali Shrivastava

**HIGHLIGHT:** Theoretically, we prove that count-sketch optimization maintains the SGD convergence rate, while gracefully reducing memory usage for large-models.

599, **TITLE:** Escaping Saddle Points with Adaptive Gradient Methods

<http://proceedings.mlr.press/v97/staib19a.html>

**AUTHORS:** Matthew Staib, Sashank Reddi, Satyen Kale, Sanjiv Kumar, Suvrit Sra

**HIGHLIGHT:** In this paper, we seek a crisp, clean and precise characterization of their behavior in nonconvex settings.

- 600, TITLE: Faster Attend-Infer-Repeat with Tractable Probabilistic Models  
<http://proceedings.mlr.press/v97/stelzner19a.html>  
AUTHORS: Karl Stelzner, Robert Peharz, Kristian Kersting  
HIGHLIGHT: In this paper, we show that the speed and robustness of learning in AIR can be considerably improved by replacing the intractable object representations with tractable probabilistic models.
- 601, TITLE: Insertion Transformer: Flexible Sequence Generation via Insertion Operations  
<http://proceedings.mlr.press/v97/stern19a.html>  
AUTHORS: Mitchell Stern, William Chan, Jamie Kiros, Jakob Uszkoreit  
HIGHLIGHT: We present the Insertion Transformer, an iterative, partially autoregressive model for sequence generation based on insertion operations.
- 602, TITLE: BERT and PALs: Projected Attention Layers for Efficient Adaptation in Multi-Task Learning  
<http://proceedings.mlr.press/v97/stickland19a.html>  
AUTHORS: Asa Cooper Stickland, Iain Murray  
HIGHLIGHT: We explore multi-task approaches that share a  $\text{\hbox{single}}$  BERT model with a small number of additional task-specific parameters.
- 603, TITLE: Learning Optimal Linear Regularizers  
<http://proceedings.mlr.press/v97/streeter19a.html>  
AUTHORS: Matthew Streeter  
HIGHLIGHT: We present algorithms for efficiently learning regularizers that improve generalization.
- 604, TITLE: CAB: Continuous Adaptive Blending for Policy Evaluation and Learning  
<http://proceedings.mlr.press/v97/su19a.html>  
AUTHORS: Yi Su, Lequn Wang, Michele Santacatterina, Thorsten Joachims  
HIGHLIGHT: In this paper, we identify a family of counterfactual estimators which subsumes most such estimators proposed to date.
- 605, TITLE: Learning Distance for Sequences by Learning a Ground Metric  
<http://proceedings.mlr.press/v97/su19b.html>  
AUTHORS: Bing Su, Ying Wu  
HIGHLIGHT: We propose to learn the distance for sequences through learning a ground Mahalanobis metric for the vectors in sequences.
- 606, TITLE: Contextual Memory Trees  
<http://proceedings.mlr.press/v97/sun19a.html>  
AUTHORS: Wen Sun, Alina Beygelzimer, Hal Daum? Iii, John Langford, Paul Mineiro  
HIGHLIGHT: We design and study a Contextual Memory Tree (CMT), a learning memory controller that inserts new memories into an experience store of unbounded size.
- 607, TITLE: Provably Efficient Imitation Learning from Observation Alone  
<http://proceedings.mlr.press/v97/sun19b.html>  
AUTHORS: Wen Sun, Anirudh Vemula, Byron Boots, Drew Bagnell  
HIGHLIGHT: We design a new model-free algorithm for ILFO, Forward Adversarial Imitation Learning (FAIL), which learns a sequence of time-dependent policies by minimizing an Integral Probability Metric between the observation distributions of the expert policy and the learner.
- 608, TITLE: Active Learning for Decision-Making from Imbalanced Observational Data  
<http://proceedings.mlr.press/v97/sundin19a.html>  
AUTHORS: Iris Sundin, Peter Schulam, Eero Siivola, Aki Vehtari, Suchi Saria, Samuel Kaski  
HIGHLIGHT: We propose to assess the decision-making reliability by estimating the ITE model's Type S error rate, which is the probability of the model inferring the sign of the treatment effect wrong.



609, TITLE: Robustly Disentangled Causal Mechanisms: Validating Deep Representations for Interventional Robustness

<http://proceedings.mlr.press/v97/suter19a.html>

AUTHORS: Raphael Suter, Djordje Miladinovic, Bernhard Schölkopf, Stefan Bauer

HIGHLIGHT: We provide a causal perspective on representation learning which covers disentanglement and domain shift robustness as special cases.

610, TITLE: Hyperbolic Disk Embeddings for Directed Acyclic Graphs

<http://proceedings.mlr.press/v97/suzuki19a.html>

AUTHORS: Ryota Suzuki, Ryusuke Takahama, Shun Onoda

HIGHLIGHT: Tackling in this problem, we develop Disk Embeddings, which is a framework for embedding DAGs into quasi-metric spaces.

611, TITLE: Accelerated Flow for Probability Distributions

<http://proceedings.mlr.press/v97/taghvaei19a.html>

AUTHORS: Amirhossein Taghvaei, Prashant Mehta

HIGHLIGHT: This paper presents a methodology and numerical algorithms for constructing accelerated gradient flows on the space of probability distributions.

612, TITLE: Equivariant Transformer Networks

<http://proceedings.mlr.press/v97/tai19a.html>

AUTHORS: Kai Sheng Tai, Peter Bailis, Gregory Valiant

HIGHLIGHT: We propose Equivariant Transformers (ETs), a family of differentiable image-to-image mappings that improve the robustness of models towards pre-defined continuous transformation groups.

613, TITLE: Making Deep Q-learning methods robust to time discretization

<http://proceedings.mlr.press/v97/tallec19a.html>

AUTHORS: Corentin Tallec, Léonard Blier, Yann Ollivier

HIGHLIGHT: In this paper, we identify sensitivity to time discretization in near continuous-time environments as a critical factor; this covers, e.g., changing the number of frames per second, or the action frequency of the controller.

614, TITLE: EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks

<http://proceedings.mlr.press/v97/tan19a.html>

AUTHORS: Mingxing Tan, Quoc Le

HIGHLIGHT: In this paper, we systematically study model scaling and identify that carefully balancing network depth, width, and resolution can lead to better performance.

615, TITLE: Hierarchical Compositional Mixtures of Variational Autoencoders

<http://proceedings.mlr.press/v97/tan19b.html>

AUTHORS: Ping Liang Tan, Robert Peharz

HIGHLIGHT: Since these problems become generally more severe in high dimensions, we propose a novel hierarchical mixture model over low-dimensional VAE experts.

616, TITLE: Mallows ranking models: maximum likelihood estimate and regeneration

<http://proceedings.mlr.press/v97/tang19a.html>

AUTHORS: Wenpin Tang

HIGHLIGHT: Motivated by the infinite top- $k$  ranking model, we propose an algorithm to select the model size  $k$  automatically.

617, TITLE: Correlated Variational Auto-Encoders

<http://proceedings.mlr.press/v97/tang19b.html>

AUTHORS: Da Tang, Dawen Liang, Tony Jebara, Nicholas Ruoizzi  
HIGHLIGHT: We propose Correlated Variational Auto-Encoders (CVAEs) that can take the correlation structure into consideration when learning latent representations with VAEs.

618, TITLE: The Variational Predictive Natural Gradient  
<http://proceedings.mlr.press/v97/tang19c.html>  
AUTHORS: Da Tang, Rajesh Ranganath  
HIGHLIGHT: To address this, we construct a new natural gradient called the Variational Predictive Natural Gradient (VPNG).

619, TITLE:  $\text{\$}\text{t\texttt{DoubleSqueeze}}\text{\$}$ : Parallel Stochastic Gradient Descent with Double-pass Error-Compensated Compression  
<http://proceedings.mlr.press/v97/tang19d.html>  
AUTHORS: Hanlin Tang, Chen Yu, Xiangru Lian, Tong Zhang, Ji Liu  
HIGHLIGHT: In this work, we provide a detailed analysis on this two-pass communication model, with error-compensated compression both on the worker nodes and on the parameter server.

620, TITLE: Adaptive Neural Trees  
<http://proceedings.mlr.press/v97/tanno19a.html>  
AUTHORS: Ryutaro Tanno, Kai Arulkumaran, Daniel Alexander, Antonio Criminisi, Aditya Nori  
HIGHLIGHT: We unite the two via adaptive neural trees (ANTs), a model that incorporates representation learning into edges, routing functions and leaf nodes of a decision tree, along with a backpropagation-based training algorithm that adaptively grows the architecture from primitive modules (e.g., convolutional layers).

621, TITLE: Variational Annealing of GANs: A Langevin Perspective  
<http://proceedings.mlr.press/v97/tao19a.html>  
AUTHORS: Chenyang Tao, Shuyang Dai, Liqun Chen, Ke Bai, Junya Chen, Chang Liu, Ruiyi Zhang, Georgiy Bobashev, Lawrence Carin Duke  
HIGHLIGHT: We highlight new insights from variational theory of diffusion processes to derive a likelihood-based regularizing scheme for GAN training, and present a novel approach to train GANs with an unnormalized distribution instead of empirical samples.

622, TITLE: Predicate Exchange: Inference with Declarative Knowledge  
<http://proceedings.mlr.press/v97/tavares19a.html>  
AUTHORS: Zenna Tavares, Rajesh Ranganath, Javier Burroni, Armando Solar-Lezama, Edgar Minasyan  
HIGHLIGHT: To support a broader class of predicates, we develop an inference procedure called predicate exchange, which softens predicates.

623, TITLE: The Natural Language of Actions  
<http://proceedings.mlr.press/v97/tennenholtz19a.html>  
AUTHORS: Guy Tennenholtz, Shie Mannor  
HIGHLIGHT: We introduce Act2Vec, a general framework for learning context-based action representation for Reinforcement Learning.

624, TITLE: Kernel Normalized Cut: a Theoretical Revisit  
<http://proceedings.mlr.press/v97/terada19a.html>  
AUTHORS: Yoshikazu Terada, Michio Yamamoto  
HIGHLIGHT: In this paper, we study the theoretical properties of clustering based on the kernel normalized cut.

625, TITLE: Action Robust Reinforcement Learning and Applications in Continuous Control  
<http://proceedings.mlr.press/v97/tessler19a.html>  
AUTHORS: Chen Tessler, Yonathan Efroni, Shie Mannor  
HIGHLIGHT: In this work we formalize two new criteria of robustness to action uncertainty.

626, TITLE: Concentration Inequalities for Conditional Value at Risk  
<http://proceedings.mlr.press/v97/thomas19a.html>  
AUTHORS: Philip Thomas, Erik Learned-Miller  
HIGHLIGHT: In this paper we derive new concentration inequalities for the conditional value at risk (CVaR) of a random variable, and compare them to the previous state of the art (Brown, 2007).

627, TITLE: Combating Label Noise in Deep Learning using Abstention  
<http://proceedings.mlr.press/v97/thulasidasan19a.html>  
AUTHORS: Sunil Thulasidasan, Tanmoy Bhattacharya, Jeff Bilmes, Gopinath Chennupati, Jamal Mohd-Yusof  
HIGHLIGHT: We introduce a novel method to combat label noise when training deep neural networks for classification.

628, TITLE: ELF OpenGo: an analysis and open reimplementation of AlphaZero  
<http://proceedings.mlr.press/v97/tian19a.html>  
AUTHORS: Yuandong Tian, Jerry Ma, Qucheng Gong, Shubho Sengupta, Zhuoyuan Chen, James Pinkerton, Larry Zitnick  
HIGHLIGHT: Toward elucidating unresolved mysteries and facilitating future research, we propose ELF OpenGo, an open-source reimplementation of the AlphaZero algorithm.

629, TITLE: Random Matrix Improved Covariance Estimation for a Large Class of Metrics  
<http://proceedings.mlr.press/v97/tiomoko19a.html>  
AUTHORS: Malik Tiomoko, Romain Couillet, Florent Bouchard, Guillaume Ginolhac  
HIGHLIGHT: Relying on recent advances in statistical estimation of covariance distances based on random matrix theory, this article proposes an improved covariance and precision matrix estimation for a wide family of metrics.

630, TITLE: Transfer of Samples in Policy Search via Multiple Importance Sampling  
<http://proceedings.mlr.press/v97/tirinizoni19a.html>  
AUTHORS: Andrea Tirinzoni, Mattia Salvini, Marcello Restelli  
HIGHLIGHT: In this paper, we consider the more complex case of reusing samples in policy search methods, in which the agent is required to transfer entire trajectories between environments with different transition models.

631, TITLE: Optimal Transport for structured data with application on graphs  
<http://proceedings.mlr.press/v97/titouan19a.html>  
AUTHORS: Vayer Titouan, Nicolas Courty, Romain Tavenard, Chapel Laetitia, R?mi Flamary  
HIGHLIGHT: This work considers the problem of computing distances between structured objects such as undirected graphs, seen as probability distributions in a specific metric space.

632, TITLE: Discovering Latent Covariance Structures for Multiple Time Series  
<http://proceedings.mlr.press/v97/tong19a.html>  
AUTHORS: Anh Tong, Jaesik Choi  
HIGHLIGHT: We present a pragmatic search algorithm which explores a larger structure space efficiently.

633, TITLE: Bayesian Generative Active Deep Learning  
<http://proceedings.mlr.press/v97/tran19a.html>  
AUTHORS: Toan Tran, Thanh-Toan Do, Ian Reid, Gustavo Carneiro  
HIGHLIGHT: In this paper, we propose a Bayesian generative active deep learning approach that combines active learning with data augmentation – we provide theoretical and empirical evidence (MNIST, CIFAR- $\{10,100\}$ , and SVHN) that our approach has more efficient training and better classification results than data augmentation and active learning.

634, TITLE: DeepNose: Using artificial neural networks to represent the space of odorants  
<http://proceedings.mlr.press/v97/tran19b.html>

AUTHORS: Ngoc Tran, Daniel Kepple, Sergey Shuvaev, Alexei Koulakov  
HIGHLIGHT: We propose that DeepNose network can extract de novo chemical features predictive of various bioactivities and can help understand the factors influencing the composition of ORs ensemble.

635, TITLE: LR-GLM: High-Dimensional Bayesian Inference Using Low-Rank Data Approximations  
<http://proceedings.mlr.press/v97/trippe19a.html>

AUTHORS: Brian Trippe, Jonathan Huggins, Raj Agrawal, Tamara Broderick  
HIGHLIGHT: We propose to reduce time and memory costs with a low-rank approximation of the data in an approach we call LR-GLM.

636, TITLE: Learning Hawkes Processes Under Synchronization Noise  
<http://proceedings.mlr.press/v97/trouleau19a.html>

AUTHORS: William Trouleau, Jalal Etesami, Matthias Grossglauser, Negar Kiyavash, Patrick Thiran  
HIGHLIGHT: We characterize the robustness of the classic maximum likelihood estimator to synchronization noise, and we introduce a new approach for learning the causal structure in the presence of noise.

637, TITLE: Homomorphic Sensing

<http://proceedings.mlr.press/v97/tsakiris19a.html>

AUTHORS: Manolis Tsakiris, Liangzu Peng  
HIGHLIGHT: In this paper we introduce an abstraction of this problem which we call "homomorphic sensing".

638, TITLE: Metropolis-Hastings Generative Adversarial Networks

<http://proceedings.mlr.press/v97/turner19a.html>

AUTHORS: Ryan Turner, Jane Hung, Eric Frank, Yunus Saatchi, Jason Yosinski  
HIGHLIGHT: We introduce the Metropolis-Hastings generative adversarial network (MH-GAN), which combines aspects of Markov chain Monte Carlo and GANs.

639, TITLE: Distributed, Egocentric Representations of Graphs for Detecting Critical Structures

<http://proceedings.mlr.press/v97/tzeng19a.html>

AUTHORS: Ruo-Chun Tzeng, Shan-Hung Wu  
HIGHLIGHT: In this paper, we propose a novel graph embedding model, called the Ego-CNNs, that employs the ego-convolutions convolutions at each layer and stacks up layers using an ego-centric way to detects precise critical structures efficiently.

640, TITLE: Sublinear Space Private Algorithms Under the Sliding Window Model

<http://proceedings.mlr.press/v97/upadhyay19a.html>

AUTHORS: Jalaj Upadhyay  
HIGHLIGHT: In this paper, we study heavy hitters in the sliding window model with window size  $\$w\$$ .

641, TITLE: Fairness without Harm: Decoupled Classifiers with Preference Guarantees

<http://proceedings.mlr.press/v97/ustun19a.html>

AUTHORS: Berk Ustun, Yang Liu, David Parkes  
HIGHLIGHT: In this work, we argue that when there is this kind of treatment disparity, then it should be in the best interest of each group.

642, TITLE: Large-Scale Sparse Kernel Canonical Correlation Analysis

<http://proceedings.mlr.press/v97/uurtio19a.html>

AUTHORS: Viivi Uurtio, Sahely Bhadra, Juho Rousu  
HIGHLIGHT: This paper presents gradKCCA, a large-scale sparse non-linear canonical correlation method.

643, TITLE: Characterization of Convex Objective Functions and Optimal Expected Convergence Rates for SGD

<http://proceedings.mlr.press/v97/van-dijk19a.html>

- AUTHORS: Marten Van Dijk, Lam Nguyen, Phuong Ha Nguyen, Dzung Phan  
HIGHLIGHT: We introduce a definitional framework and theory that defines and characterizes a core property, called curvature, of convex objective functions.
- 644, TITLE: Composing Value Functions in Reinforcement Learning  
<http://proceedings.mlr.press/v97/van-niekerk19a.html>  
AUTHORS: Benjamin Van Niekerk, Steven James, Adam Earle, Benjamin Rosman  
HIGHLIGHT: Under the assumption of deterministic dynamics, we prove that optimal value function composition can be achieved in entropy-regularised reinforcement learning (RL), and extend this result to the standard RL setting.
- 645, TITLE: Model Comparison for Semantic Grouping  
<http://proceedings.mlr.press/v97/vargas19a.html>  
AUTHORS: Francisco Vargas, Kamen Brestnichki, Nils Hammerla  
HIGHLIGHT: We introduce a probabilistic framework for quantifying the semantic similarity between two groups of embeddings.
- 646, TITLE: Learning Dependency Structures for Weak Supervision Models  
<http://proceedings.mlr.press/v97/varma19a.html>  
AUTHORS: Paroma Varma, Frederic Sala, Ann He, Alexander Ratner, Christopher Re  
HIGHLIGHT: We focus on a robust PCA-based algorithm for learning these dependency structures, establish improved theoretical recovery rates, and outperform existing methods on various real-world tasks.
- 647, TITLE: Probabilistic Neural Symbolic Models for Interpretable Visual Question Answering  
<http://proceedings.mlr.press/v97/vedantam19a.html>  
AUTHORS: Ramakrishna Vedantam, Karan Desai, Stefan Lee, Marcus Rohrbach, Dhruv Batra, Devi Parikh  
HIGHLIGHT: We propose a new class of probabilistic neural-symbolic models, that have symbolic functional programs as a latent, stochastic variable.
- 648, TITLE: Manifold Mixup: Better Representations by Interpolating Hidden States  
<http://proceedings.mlr.press/v97/verma19a.html>  
AUTHORS: Vikas Verma, Alex Lamb, Christopher Beckham, Amir Najafi, Ioannis Mitliagkas, David Lopez-Paz, Yoshua Bengio  
HIGHLIGHT: To address these issues, we propose  $\text{manifoldmixup}\{\}$ , a simple regularizer that encourages neural networks to predict less confidently on interpolations of hidden representations.
- 649, TITLE: Maximum Likelihood Estimation for Learning Populations of Parameters  
<http://proceedings.mlr.press/v97/vinayak19a.html>  
AUTHORS: Ramya Korlakai Vinayak, Weihao Kong, Gregory Valiant, Sham Kakade  
HIGHLIGHT: After observing the outcomes of  $t$  independent Bernoulli trials, i.e.,  $X_i \sim \text{Binomial}(t, p_i)$  per individual, our objective is to accurately estimate  $\mathbb{P}^{\star}$  in the sparse regime, namely when  $t \ll N$ .
- 650, TITLE: Understanding Priors in Bayesian Neural Networks at the Unit Level  
<http://proceedings.mlr.press/v97/vladimirova19a.html>  
AUTHORS: Mariia Vladimirova, Jakob Verbeek, Pablo Mesejo, Julyan Arbel  
HIGHLIGHT: We investigate deep Bayesian neural networks with Gaussian priors on the weights and a class of ReLU-like nonlinearities.
- 651, TITLE: On the Design of Estimators for Bandit Off-Policy Evaluation  
<http://proceedings.mlr.press/v97/vlassis19a.html>  
AUTHORS: Nikos Vlassis, Aurelien Bibaut, Maria Dimakopoulou, Tony Jebara  
HIGHLIGHT: We present our main results in the context of multi-armed bandits, and we propose a simple design for contextual bandits that gives rise to an estimator that is shown to perform well in multi-class cost-sensitive classification datasets.

- 652, TITLE: Learning to select for a predefined ranking  
<http://proceedings.mlr.press/v97/vorobev19a.html>  
AUTHORS: Aleksandr Vorobev, Aleksei Ustimenko, Gleb Gusev, Pavel Serdyukov  
HIGHLIGHT: In this paper, we formulate a novel problem of learning to select a set of items maximizing the quality of their ordered list, where the order is predefined by some explicit rule.
- 653, TITLE: On the Limitations of Representing Functions on Sets  
<http://proceedings.mlr.press/v97/wagstaff19a.html>  
AUTHORS: Edward Wagstaff, Fabian Fuchs, Martin Engelcke, Ingmar Posner, Michael A. Osborne  
HIGHLIGHT: Motivated by this observation, we prove that an implementation of this model via continuous mappings (as provided by e.g. neural networks or Gaussian processes) actually imposes a constraint on the dimensionality of the latent space.
- 654, TITLE: Graph Convolutional Gaussian Processes  
<http://proceedings.mlr.press/v97/walker19a.html>  
AUTHORS: Ian Walker, Ben Glocker  
HIGHLIGHT: We propose a novel Bayesian nonparametric method to learn translation-invariant relationships on non-Euclidean domains.
- 655, TITLE: Gaining Free or Low-Cost Interpretability with Interpretable Partial Substitute  
<http://proceedings.mlr.press/v97/wang19a.html>  
AUTHORS: Tong Wang  
HIGHLIGHT: Under this framework, we develop a Hybrid Rule Sets (HyRS) model that uses decision rules to capture the subspace of data where the rules are as accurate or almost as accurate as the black-box provided.
- 656, TITLE: Convolutional Poisson Gamma Belief Network  
<http://proceedings.mlr.press/v97/wang19b.html>  
AUTHORS: Chaojie Wang, Bo Chen, Sucheng Xiao, Mingyuan Zhou  
HIGHLIGHT: In this paper, we propose convolutional Poisson factor analysis (CPFA) that directly operates on a lossless representation that processes the words in each document as a sequence of high-dimensional one-hot vectors.
- 657, TITLE: Differentially Private Empirical Risk Minimization with Non-convex Loss Functions  
<http://proceedings.mlr.press/v97/wang19c.html>  
AUTHORS: Di Wang, Changyou Chen, Jinhui Xu  
HIGHLIGHT: We study the problem of Empirical Risk Minimization (ERM) with (smooth) non-convex loss functions under the differential-privacy (DP) model.
- 658, TITLE: Random Expert Distillation: Imitation Learning via Expert Policy Support Estimation  
<http://proceedings.mlr.press/v97/wang19d.html>  
AUTHORS: Ruohan Wang, Carlo Ciliberto, Pierluigi Vito Amadori, Yiannis Demiris  
HIGHLIGHT: We propose a new framework for imitation learning by estimating the support of the expert policy to compute a fixed reward function, which allows us to re-frame imitation learning within the standard reinforcement learning setting.
- 659, TITLE: SATNet: Bridging deep learning and logical reasoning using a differentiable satisfiability solver  
<http://proceedings.mlr.press/v97/wang19e.html>  
AUTHORS: Po-Wei Wang, Priya Donti, Bryan Wilder, Zico Kolter  
HIGHLIGHT: In this paper, we propose a new direction toward this goal by introducing a differentiable (smoothed) maximum satisfiability (MAXSAT) solver that can be integrated into the loop of larger deep learning systems.
- 660, TITLE: Improving Neural Language Modeling via Adversarial Training  
<http://proceedings.mlr.press/v97/wang19f.html>

- AUTHORS: Dilin Wang, Chengyue Gong, Qiang Liu  
HIGHLIGHT: In this paper, we present a simple yet highly effective adversarial training mechanism for regularizing neural language models.
- 661, TITLE: EigenDamage: Structured Pruning in the Kronecker-Factored Eigenbasis  
<http://proceedings.mlr.press/v97/wang19g.html>  
AUTHORS: Chaoqi Wang, Roger Grosse, Sanja Fidler, Guodong Zhang  
HIGHLIGHT: In particular, we highlight that the improvements are especially significant for more challenging datasets and networks.
- 662, TITLE: Nonlinear Stein Variational Gradient Descent for Learning Diversified Mixture Models  
<http://proceedings.mlr.press/v97/wang19h.html>  
AUTHORS: Dilin Wang, Qiang Liu  
HIGHLIGHT: In this work, we present a variational approach for diversity-promoting learning, which leverages the entropy functional as a natural mechanism for enforcing diversity.
- 663, TITLE: On the Convergence and Robustness of Adversarial Training  
<http://proceedings.mlr.press/v97/wang19i.html>  
AUTHORS: Yisen Wang, Xingjun Ma, James Bailey, Jinfeng Yi, Bowen Zhou, Quanquan Gu  
HIGHLIGHT: In this paper, we propose such a criterion, namely First-Order Stationary Condition for constrained optimization (FOSC), to quantitatively evaluate the convergence quality of adversarial examples found in the inner maximization.
- 664, TITLE: State-Regularized Recurrent Neural Networks  
<http://proceedings.mlr.press/v97/wang19j.html>  
AUTHORS: Cheng Wang, Mathias Niepert  
HIGHLIGHT: We aim to address both shortcomings with a class of recurrent networks that use a stochastic state transition mechanism between cell applications.
- 665, TITLE: Deep Factors for Forecasting  
<http://proceedings.mlr.press/v97/wang19k.html>  
AUTHORS: Yuyang Wang, Alex Smola, Danielle Maddix, Jan Gasthaus, Dean Foster, Tim Januschowski  
HIGHLIGHT: In this paper, we propose a hybrid model that incorporates the benefits of both approaches.
- 666, TITLE: Repairing without Retraining: Avoiding Disparate Impact with Counterfactual Distributions  
<http://proceedings.mlr.press/v97/wang19l.html>  
AUTHORS: Hao Wang, Berk Ustun, Flavio Calmon  
HIGHLIGHT: In this paper, we exploit this fact to reduce the disparate impact of a fixed classification model over a population of interest.
- 667, TITLE: On Sparse Linear Regression in the Local Differential Privacy Model  
<http://proceedings.mlr.press/v97/wang19m.html>  
AUTHORS: Di Wang, Jinhui Xu  
HIGHLIGHT: In this paper, we study the sparse linear regression problem under the Local Differential Privacy (LDP) model.
- 668, TITLE: Doubly Robust Joint Learning for Recommendation on Data Missing Not at Random  
<http://proceedings.mlr.press/v97/wang19n.html>  
AUTHORS: Xiaojie Wang, Rui Zhang, Yu Sun, Jianzhong Qi  
HIGHLIGHT: To achieve good performance guarantees, based on this estimator, we propose joint learning of rating prediction and error imputation, which outperforms the state-of-the-art approaches on four real-world datasets.
- 669, TITLE: On the Generalization Gap in Reparameterizable Reinforcement Learning

<http://proceedings.mlr.press/v97/wang19o.html>

AUTHORS: Huan Wang, Stephan Zheng, Caiming Xiong, Richard Socher

HIGHLIGHT: We focus on the special class of reparameterizable RL problems, where the trajectory distribution can be decomposed using the reparametrization trick.

670, TITLE: Bias Also Matters: Bias Attribution for Deep Neural Network Explanation

<http://proceedings.mlr.press/v97/wang19p.html>

AUTHORS: Shengjie Wang, Tianyi Zhou, Jeff Bilmes

HIGHLIGHT: In this paper, we observe that since the bias in a DNN also has a non-negligible contribution to the correctness of predictions, it can also play a significant role in understanding DNN behavior.

671, TITLE: Jumpout : Improved Dropout for Deep Neural Networks with ReLUs

<http://proceedings.mlr.press/v97/wang19q.html>

AUTHORS: Shengjie Wang, Tianyi Zhou, Jeff Bilmes

HIGHLIGHT: We discuss three novel insights about dropout for DNNs with ReLUs: 1) dropout encourages each local linear piece of a DNN to be trained on data points from nearby regions; 2) the same dropout rate results in different (effective) deactivation rates for layers with different portions of ReLU-deactivated neurons; and 3) the rescaling factor of dropout causes a normalization inconsistency between training and test when used together with batch normalization.

672, TITLE: AdaGrad stepsizes: sharp convergence over nonconvex landscapes

<http://proceedings.mlr.press/v97/ward19a.html>

AUTHORS: Rachel Ward, Xiaoxia Wu, Leon Bottou

HIGHLIGHT: We bridge this gap by providing strong theoretical guarantees for the convergence of AdaGrad over smooth, nonconvex landscapes.

673, TITLE: Generalized Linear Rule Models

<http://proceedings.mlr.press/v97/wei19a.html>

AUTHORS: Dennis Wei, Sanjeeb Dash, Tian Gao, Oktay Gunluk

HIGHLIGHT: This paper considers generalized linear models using rule-based features, also referred to as rule ensembles, for regression and probabilistic classification.

674, TITLE: On the statistical rate of nonlinear recovery in generative models with heavy-tailed data

<http://proceedings.mlr.press/v97/wei19b.html>

AUTHORS: Xiaohan Wei, Zhuoran Yang, Zhaoran Wang

HIGHLIGHT: In this paper, we make a step towards such a direction by considering the scenario where the measurements are non-Gaussian, subject to possibly unknown nonlinear transformations and the responses are heavy-tailed.

675, TITLE: CapsAndRuns: An Improved Method for Approximately Optimal Algorithm Configuration

<http://proceedings.mlr.press/v97/weisz19a.html>

AUTHORS: Gellert Weisz, Andras Gyorgy, Csaba Szepesvari

HIGHLIGHT: In this paper we present a new algorithm, CapsAndRuns, which finds a near-optimal configuration while using time that scales (in a problem dependent way) with the optimal expected capped runtime, significantly strengthening previous results which could only guarantee a bound that scaled with the potentially much larger optimal expected uncapped runtime.

676, TITLE: Non-Monotonic Sequential Text Generation

<http://proceedings.mlr.press/v97/welleck19a.html>

AUTHORS: Sean Welleck, Kiant? Brantley, Hal Daum? Iii, Kyunghyun Cho

HIGHLIGHT: In this work, we propose a framework for training models of text generation that operate in non-monotonic orders; the model directly learns good orders, without any additional annotation.

677, TITLE: PROVEN: Verifying Robustness of Neural Networks with a Probabilistic Approach



<http://proceedings.mlr.press/v97/weng19a.html>

AUTHORS: Lily Weng, Pin-Yu Chen, Lam Nguyen, Mark Squillante, Akhilan Boopathy, Ivan Oseledets, Luca Daniel

HIGHLIGHT: We propose a novel framework PROVEN to probabilistically verify neural network's robustness with statistical guarantees.

678, TITLE: Learning deep kernels for exponential family densities

<http://proceedings.mlr.press/v97/wenliang19a.html>

AUTHORS: Li Wenliang, Dougal Sutherland, Heiko Strathmann, Arthur Gretton

HIGHLIGHT: We provide a scheme for learning a kernel parameterized by a deep network, which can find complex location-dependent local features of the data geometry.

679, TITLE: Improving Model Selection by Employing the Test Data

<http://proceedings.mlr.press/v97/westphal19a.html>

AUTHORS: Max Westphal, Werner Brannath

HIGHLIGHT: We investigate the properties of novel evaluation strategies, namely when the final model is selected based on empirical performances on the test data.

680, TITLE: Automatic Classifiers as Scientific Instruments: One Step Further Away from Ground-Truth

<http://proceedings.mlr.press/v97/whitehill19a.html>

AUTHORS: Jacob Whitehill, Anand Ramakrishnan

HIGHLIGHT: We examine how the accuracy of  $d$ , as quantified by the correlation  $\rho$  of  $d$ 's outputs with the ground-truth construct  $U$ , impacts the estimated correlation between  $U$  (e.g., stress) and some other phenomenon  $V$  (e.g., academic performance).

681, TITLE: Moment-Based Variational Inference for Markov Jump Processes

<http://proceedings.mlr.press/v97/wildner19a.html>

AUTHORS: Christian Wildner, Heinz Koepl

HIGHLIGHT: We propose moment-based variational inference as a flexible framework for approximate smoothing of latent Markov jump processes.

682, TITLE: End-to-End Probabilistic Inference for Nonstationary Audio Analysis

<http://proceedings.mlr.press/v97/wilkinson19a.html>

AUTHORS: William Wilkinson, Michael Andersen, Joshua D. Reiss, Dan Stowell, Arno Solin

HIGHLIGHT: We show how time-frequency analysis and nonnegative matrix factorisation can be jointly formulated as a spectral mixture Gaussian process model with nonstationary priors over the amplitude variance parameters.

683, TITLE: Fairness risk measures

<http://proceedings.mlr.press/v97/williamson19a.html>

AUTHORS: Robert Williamson, Aditya Menon

HIGHLIGHT: In this paper, we propose a new definition of fairness that generalises some existing proposals, while allowing for generic sensitive features and resulting in a convex objective.

684, TITLE: Partially Exchangeable Networks and Architectures for Learning Summary Statistics in Approximate Bayesian Computation

<http://proceedings.mlr.press/v97/wiqvist19a.html>

AUTHORS: Samuel Wiqvist, Pierre-Alexandre Mattei, Umberto Picchini, Jes Frelsen

HIGHLIGHT: We present a novel family of deep neural architectures, named partially exchangeable networks (PENs) that leverage probabilistic symmetries.

685, TITLE: Wasserstein Adversarial Examples via Projected Sinkhorn Iterations

<http://proceedings.mlr.press/v97/wong19a.html>

AUTHORS: Eric Wong, Frank Schmidt, Zico Kolter

**HIGHLIGHT:** In this paper, we propose a new threat model for adversarial attacks based on the Wasserstein distance.

686, **TITLE:** Imitation Learning from Imperfect Demonstration

<http://proceedings.mlr.press/v97/wu19a.html>

**AUTHORS:** Yueh-Hua Wu, Nontawat Charoenphakdee, Han Bao, Voot Tangkaratt, Masashi Sugiyama

**HIGHLIGHT:** To effectively learn from imperfect demonstrations, we propose a novel approach that utilizes confidence scores, which describe the quality of demonstrations.

687, **TITLE:** Learning a Compressed Sensing Measurement Matrix via Gradient Unrolling

<http://proceedings.mlr.press/v97/wu19b.html>

**AUTHORS:** Shanshan Wu, Alex Dimakis, Sujay Sanghavi, Felix Yu, Daniel Holtmann-Rice, Dmitry Storcheus, Afshin Rostamizadeh, Sanjiv Kumar

**HIGHLIGHT:** In this paper we present a new method to learn linear encoders that adapt to data, while still performing well with the widely used  $\ell_1$  decoder.

688, **TITLE:** Heterogeneous Model Reuse via Optimizing Multiparty Multiclass Margin

<http://proceedings.mlr.press/v97/wu19c.html>

**AUTHORS:** Xi-Zhu Wu, Song Liu, Zhi-Hua Zhou

**HIGHLIGHT:** In this paper, we define a multiparty multiclass margin to measure the global behavior of a set of heterogeneous local models, and propose a general learning method called HMR (Heterogeneous Model Reuse) to optimize the margin.

689, **TITLE:** Deep Compressed Sensing

<http://proceedings.mlr.press/v97/wu19d.html>

**AUTHORS:** Yan Wu, Mihaela Rosca, Timothy Lillicrap

**HIGHLIGHT:** Here we propose a novel framework that significantly improves both the performance and speed of signal recovery by jointly training a generator and the optimisation process for reconstruction via meta-learning.

690, **TITLE:** Simplifying Graph Convolutional Networks

<http://proceedings.mlr.press/v97/wu19e.html>

**AUTHORS:** Felix Wu, Amauri Souza, Tianyi Zhang, Christopher Fifty, Tao Yu, Kilian Weinberger

**HIGHLIGHT:** In this paper, we reduce this excess complexity through successively removing nonlinearities and collapsing weight matrices between consecutive layers.

691, **TITLE:** Domain Adaptation with Asymmetrically-Relaxed Distribution Alignment

<http://proceedings.mlr.press/v97/wu19f.html>

**AUTHORS:** Yifan Wu, Ezra Winston, Divyansh Kaushik, Zachary Lipton

**HIGHLIGHT:** We propose asymmetrically-relaxed distribution alignment, a new approach that overcomes some limitations of standard domain-adversarial algorithms.

692, **TITLE:** On Scalable and Efficient Computation of Large Scale Optimal Transport

<http://proceedings.mlr.press/v97/xie19a.html>

**AUTHORS:** Yujia Xie, Minshuo Chen, Haoming Jiang, Tuo Zhao, Hongyuan Zha

**HIGHLIGHT:** To address the scalability issue, we propose an implicit generative learning-based framework called SPOT (Scalable Push-forward of Optimal Transport).

693, **TITLE:** Zeno: Distributed Stochastic Gradient Descent with Suspicion-based Fault-tolerance

<http://proceedings.mlr.press/v97/xie19b.html>

**AUTHORS:** Cong Xie, Sanmi Koyejo, Indranil Gupta

**HIGHLIGHT:** We present Zeno, a technique to make distributed machine learning, particularly Stochastic Gradient Descent (SGD), tolerant to an arbitrary number of faulty workers.

694, TITLE: Differentiable Linearized ADMM  
<http://proceedings.mlr.press/v97/xie19c.html>  
AUTHORS: Xingyu Xie, Jianlong Wu, Guangcan Liu, Zhisheng Zhong, Zhouchen Lin  
HIGHLIGHT: In this paper, we propose Differentiable Linearized ADMM (D-LADMM) for solving the problems with linear constraints.

695, TITLE: Calibrated Approximate Bayesian Inference  
<http://proceedings.mlr.press/v97/xing19a.html>  
AUTHORS: Hanwen Xing, Geoff Nicholls, Jeong Lee  
HIGHLIGHT: We give a general purpose computational framework for estimating the bias in coverage resulting from making approximations in Bayesian inference.

696, TITLE: Power k-Means Clustering  
<http://proceedings.mlr.press/v97/xu19a.html>  
AUTHORS: Jason Xu, Kenneth Lange  
HIGHLIGHT: This paper explores an alternative to Lloyd's algorithm that retains its simplicity and mitigates its tendency to get trapped by local minima.

697, TITLE: Gromov-Wasserstein Learning for Graph Matching and Node Embedding  
<http://proceedings.mlr.press/v97/xu19b.html>  
AUTHORS: Hongteng Xu, Dixin Luo, Hongyuan Zha, Lawrence Carin Duke  
HIGHLIGHT: We apply the proposed method to matching problems in real-world networks, and demonstrate its superior performance compared to alternative approaches.

698, TITLE: Stochastic Optimization for DC Functions and Non-smooth Non-convex Regularizers with Non-asymptotic Convergence  
<http://proceedings.mlr.press/v97/xu19c.html>  
AUTHORS: Yi Xu, Qi Qi, Qihang Lin, Rong Jin, Tianbao Yang  
HIGHLIGHT: In this paper, we propose new stochastic optimization algorithms and study their first-order convergence theories for solving a broad family of DC functions.

699, TITLE: Learning a Prior over Intent via Meta-Inverse Reinforcement Learning  
<http://proceedings.mlr.press/v97/xu19d.html>  
AUTHORS: Kelvin Xu, Ellis Ratner, Anca Dragan, Sergey Levine, Chelsea Finn  
HIGHLIGHT: In this work, we exploit the insight that demonstrations from other tasks can be used to constrain the set of possible reward functions by learning a "prior" that is specifically optimized for the ability to infer expressive reward functions from limited numbers of demonstrations.

700, TITLE: Variational Russian Roulette for Deep Bayesian Nonparametrics  
<http://proceedings.mlr.press/v97/xu19e.html>  
AUTHORS: Kai Xu, Akash Srivastava, Charles Sutton  
HIGHLIGHT: Instead, we propose a new variational approximation, based on a method from statistical physics called Russian roulette sampling.

701, TITLE: Supervised Hierarchical Clustering with Exponential Linkage  
<http://proceedings.mlr.press/v97/yadav19a.html>  
AUTHORS: Nishant Yadav, Ari Kobren, Nicholas Monath, Andrew McCallum  
HIGHLIGHT: In this paper, we introduce a method for training the dissimilarity function in a way that is tightly coupled with hierarchical clustering, in particular single linkage.

702, TITLE: Learning to Prove Theorems via Interacting with Proof Assistants  
<http://proceedings.mlr.press/v97/yang19a.html>  
AUTHORS: Kaiyu Yang, Jia Deng

**HIGHLIGHT:** In this paper, we study the problem of using machine learning to automate the interaction with proof assistants.

703, **TITLE:** Sample-Optimal Parametric Q-Learning Using Linearly Additive Features  
<http://proceedings.mlr.press/v97/yang19b.html>  
**AUTHORS:** Lin Yang, Mengdi Wang  
**HIGHLIGHT:** We propose a parametric Q-learning algorithm that finds an approximate-optimal policy using a sample size proportional to the feature dimension  $d$  and invariant with respect to the size of the state space.

704, **TITLE:** LegoNet: Efficient Convolutional Neural Networks with Lego Filters  
<http://proceedings.mlr.press/v97/yang19c.html>  
**AUTHORS:** Zhaohui Yang, Yunhe Wang, Chuanjian Liu, Hanting Chen, Chunjing Xu, Boxin Shi, Chao Xu, Chang Xu  
**HIGHLIGHT:** This paper aims to build efficient convolutional neural networks using a set of Lego filters.

705, **TITLE:** SWALP : Stochastic Weight Averaging in Low Precision Training  
<http://proceedings.mlr.press/v97/yang19d.html>  
**AUTHORS:** Guandao Yang, Tianyi Zhang, Polina Kirichenko, Junwen Bai, Andrew Gordon Wilson, Chris De Sa  
**HIGHLIGHT:** This paper proposes SWALP, an approach to low precision training that averages low-precision SGD iterates with a modified learning rate schedule.

706, **TITLE:** ME-Net: Towards Effective Adversarial Robustness with Matrix Estimation  
<http://proceedings.mlr.press/v97/yang19e.html>  
**AUTHORS:** Yuzhe Yang, Guo Zhang, Zhi Xu, Dina Katabi  
**HIGHLIGHT:** This paper proposes ME-Net, a defense method that leverages matrix estimation (ME).

707, **TITLE:** Efficient Nonconvex Regularized Tensor Completion with Structure-aware Proximal Iterations  
<http://proceedings.mlr.press/v97/yao19a.html>  
**AUTHORS:** Quanming Yao, James Tin-Yau Kwok, Bo Han  
**HIGHLIGHT:** In this paper, we extend this to the more challenging problem of low-rank tensor completion.

708, **TITLE:** Hierarchically Structured Meta-learning  
<http://proceedings.mlr.press/v97/yao19b.html>  
**AUTHORS:** Huaxiu Yao, Ying Wei, Junzhou Huang, Zhenhui Li  
**HIGHLIGHT:** In this paper, based on gradient-based meta-learning, we propose a hierarchically structured meta-learning (HSML) algorithm that explicitly tailors the transferable knowledge to different clusters of tasks.

709, **TITLE:** Tight Kernel Query Complexity of Kernel Ridge Regression and Kernel  $k$ -means Clustering  
<http://proceedings.mlr.press/v97/yasuda19a.html>  
**AUTHORS:** Taisuke Yasuda, David Woodruff, Manuel Fernandez  
**HIGHLIGHT:** In this work, we present nearly tight lower bounds on the number of kernel evaluations required to approximately solve kernel ridge regression (KRR) and kernel  $k$ -means clustering (KKMC) on  $n$  input points.

710, **TITLE:** Understanding Geometry of Encoder-Decoder CNNs  
<http://proceedings.mlr.press/v97/ye19a.html>  
**AUTHORS:** Jong Chul Ye, Woon Kyoung Sung  
**HIGHLIGHT:** Inspired by recent theoretical understanding on generalizability, expressivity and optimization landscape of neural networks, as well as the theory of convolutional framelets, here we provide a unified theoretical framework that leads to a better understanding of geometry of encoder-decoder CNNs.

711, **TITLE:** Defending Against Saddle Point Attack in Byzantine-Robust Distributed Learning  
<http://proceedings.mlr.press/v97/yin19a.html>

AUTHORS: Dong Yin, Yudong Chen, Ramchandran Kannan, Peter Bartlett  
HIGHLIGHT: As a by-product, we give a simpler algorithm and analysis for escaping saddle points in the usual non-Byzantine setting.

712, TITLE: Rademacher Complexity for Adversarially Robust Generalization  
<http://proceedings.mlr.press/v97/yin19b.html>

AUTHORS: Dong Yin, Ramchandran Kannan, Peter Bartlett  
HIGHLIGHT: In this paper, we focus on  $\ell_\infty$  attacks, and study the adversarially robust generalization problem through the lens of Rademacher complexity.

713, TITLE: ARSM: Augment-REINFORCE-Swap-Merge Estimator for Gradient Backpropagation Through Categorical Variables  
<http://proceedings.mlr.press/v97/yin19c.html>

AUTHORS: Mingzhang Yin, Yuguang Yue, Mingyuan Zhou  
HIGHLIGHT: To address the challenge of backpropagating the gradient through categorical variables, we propose the augment-REINFORCE-swap-merge (ARSM) gradient estimator that is unbiased and has low variance.

714, TITLE: NAS-Bench-101: Towards Reproducible Neural Architecture Search  
<http://proceedings.mlr.press/v97/ying19a.html>

AUTHORS: Chris Ying, Aaron Klein, Eric Christiansen, Esteban Real, Kevin Murphy, Frank Hutter  
HIGHLIGHT: We aim to ameliorate these problems by introducing NAS-Bench-101, the first public architecture dataset for NAS research.

715, TITLE: TapNet: Neural Network Augmented with Task-Adaptive Projection for Few-Shot Learning  
<http://proceedings.mlr.press/v97/yoon19a.html>

AUTHORS: Sung Whan Yoon, Jun Seo, Jaekyun Moon  
HIGHLIGHT: We propose TapNets, neural networks augmented with task-adaptive projection for improved few-shot learning.

716, TITLE: Towards Accurate Model Selection in Deep Unsupervised Domain Adaptation  
<http://proceedings.mlr.press/v97/you19a.html>

AUTHORS: Kaichao You, Ximei Wang, Mingsheng Long, Michael Jordan  
HIGHLIGHT: To this end, we propose Deep Embedded Validation (DEV), which embeds adapted feature representation into the validation procedure to obtain unbiased estimation of the target risk with bounded variance.

717, TITLE: Position-aware Graph Neural Networks  
<http://proceedings.mlr.press/v97/you19b.html>

AUTHORS: Jiaxuan You, Rex Ying, Jure Leskovec  
HIGHLIGHT: Here we propose Position-aware Graph Neural Networks (P-GNNs), a new class of GNNs for computing position-aware node embeddings.

718, TITLE: Learning Neurosymbolic Generative Models via Program Synthesis  
<http://proceedings.mlr.press/v97/young19a.html>

AUTHORS: Halley Young, Osbert Bastani, Mayur Naik  
HIGHLIGHT: We propose to address this problem by incorporating programs representing global structure into generative models—e.g., a 2D for-loop may represent a repeating pattern of windows—along with a framework for learning these models by leveraging program synthesis to obtain training data.

719, TITLE: DAG-GNN: DAG Structure Learning with Graph Neural Networks  
<http://proceedings.mlr.press/v97/you19a.html>

AUTHORS: Yue Yu, Jie Chen, Tian Gao, Mo Yu  
HIGHLIGHT: Motivated by the widespread success of deep learning that is capable of capturing complex nonlinear mappings, in this work we propose a deep generative model and apply a variant of the structural constraint to learn the DAG.

720, TITLE: How does Disagreement Help Generalization against Label Corruption?  
<http://proceedings.mlr.press/v97/yu19b.html>  
AUTHORS: Xingrui Yu, Bo Han, Jiangchao Yao, Gang Niu, Ivor Tsang, Masashi Sugiyama  
HIGHLIGHT: To tackle this issue, we propose a robust learning paradigm called Co-teaching+, which bridges the "Update by Disagreement" strategy with the original Co-teaching.

721, TITLE: On the Computation and Communication Complexity of Parallel SGD with Dynamic Batch Sizes for Stochastic Non-Convex Optimization  
<http://proceedings.mlr.press/v97/yu19c.html>  
AUTHORS: Hao Yu, Rong Jin  
HIGHLIGHT: For general stochastic non-convex optimization, we propose a Catalyst-like algorithm to achieve the fastest known  $\mathcal{O}(1/\sqrt{NT})$  convergence with only  $\mathcal{O}(\sqrt{NT} \log(\frac{T}{N}))$  communication rounds.

722, TITLE: On the Linear Speedup Analysis of Communication Efficient Momentum SGD for Distributed Non-Convex Optimization  
<http://proceedings.mlr.press/v97/yu19d.html>  
AUTHORS: Hao Yu, Rong Jin, Sen Yang  
HIGHLIGHT: This paper fills the gap by considering a distributed communication efficient momentum SGD method and proving its linear speedup property.

723, TITLE: Multi-Agent Adversarial Inverse Reinforcement Learning  
<http://proceedings.mlr.press/v97/yu19e.html>  
AUTHORS: Lantao Yu, Jiaming Song, Stefano Ermon  
HIGHLIGHT: In this paper, we propose MA-AIRL, a new framework for multi-agent inverse reinforcement learning, which is effective and scalable for Markov games with high-dimensional state-action space and unknown dynamics.

724, TITLE: Distributed Learning over Unreliable Networks  
<http://proceedings.mlr.press/v97/yu19f.html>  
AUTHORS: Chen Yu, Hanlin Tang, Cedric Renggli, Simon Kassing, Ankit Singla, Dan Alistarh, Ce Zhang, Ji Liu  
HIGHLIGHT: In this paper, we connect these two trends, and consider the following question: Can we design machine learning systems that are tolerant to network unreliability during training?

725, TITLE: Online Adaptive Principal Component Analysis and Its extensions  
<http://proceedings.mlr.press/v97/yuan19a.html>  
AUTHORS: Jianjun Yuan, Andrew Lamperski  
HIGHLIGHT: We propose algorithms for online principal component analysis (PCA) and variance minimization for adaptive settings.

726, TITLE: Generative Modeling of Infinite Occluded Objects for Compositional Scene Representation  
<http://proceedings.mlr.press/v97/yuan19b.html>  
AUTHORS: Jinyang Yuan, Bin Li, Xiangyang Xue  
HIGHLIGHT: We present a deep generative model which explicitly models object occlusions for compositional scene representation.

727, TITLE: Differential Inclusions for Modeling Nonsmooth ADMM Variants: A Continuous Limit Theory  
<http://proceedings.mlr.press/v97/yuan19c.html>  
AUTHORS: Huizhuo Yuan, Yuren Zhou, Chris Junchi Li, Qingyun Sun  
HIGHLIGHT: In this paper, we analyze some well-known and widely used ADMM variants for nonsmooth optimization problems using tools of differential inclusions.

728, TITLE: Trimming the  $\ell_1$  Regularizer: Statistical Analysis, Optimization, and Applications to Deep Learning

<http://proceedings.mlr.press/v97/yun19a.html>

AUTHORS: Jihun Yun, Peng Zheng, Eunho Yang, Aurelie Lozano, Aleksandr Aravkin

HIGHLIGHT: We present the first statistical analyses for M-estimation, and characterize support recovery,  $\ell_\infty$  and  $\ell_2$  error of the trimmed  $\ell_1$  estimates as a function of the trimming parameter  $h$ .

729, TITLE: Bayesian Nonparametric Federated Learning of Neural Networks

<http://proceedings.mlr.press/v97/yurochkin19a.html>

AUTHORS: Mikhail Yurochkin, Mayank Agarwal, Soumya Ghosh, Kristjan Greenewald, Nghia Hoang, Yasaman Khazaeni

HIGHLIGHT: We develop a Bayesian nonparametric framework for federated learning with neural networks.

730, TITLE: Dirichlet Simplex Nest and Geometric Inference

<http://proceedings.mlr.press/v97/yurochkin19b.html>

AUTHORS: Mikhail Yurochkin, Aritra Guha, Yuekai Sun, Xuanlong Nguyen

HIGHLIGHT: We propose Dirichlet Simplex Nest, a class of probabilistic models suitable for a variety of data types, and develop fast and provably accurate inference algorithms by accounting for the model's convex geometry and low dimensional simplicial structure.

731, TITLE: A Conditional-Gradient-Based Augmented Lagrangian Framework

<http://proceedings.mlr.press/v97/yurtsever19a.html>

AUTHORS: Alp Yurtsever, Olivier Fercoq, Volkan Cevher

HIGHLIGHT: To this end, we propose a new conditional gradient method, based on a unified treatment of smoothing and augmented Lagrangian frameworks.

732, TITLE: Conditional Gradient Methods via Stochastic Path-Integrated Differential Estimator

<http://proceedings.mlr.press/v97/yurtsever19b.html>

AUTHORS: Alp Yurtsever, Suvrit Sra, Volkan Cevher

HIGHLIGHT: We propose a class of variance-reduced stochastic conditional gradient methods.

733, TITLE: Context-Aware Zero-Shot Learning for Object Recognition

<http://proceedings.mlr.press/v97/zablocki19a.html>

AUTHORS: Eloi Zablocki, Patrick Bordes, Laure Soulier, Benjamin Piwowarski, Patrick Gallinari

HIGHLIGHT: Following the intuitive principle that objects tend to be found in certain contexts but not others, we propose a new and challenging approach, context-aware ZSL, that leverages semantic representations in a new way to model the conditional likelihood of an object to appear in a given context.

734, TITLE: Tighter Problem-Dependent Regret Bounds in Reinforcement Learning without Domain Knowledge using Value Function Bounds

<http://proceedings.mlr.press/v97/zanette19a.html>

AUTHORS: Andrea Zanette, Emma Brunskill

HIGHLIGHT: As a step towards this we derive an algorithm and analysis for finite horizon discrete MDPs with state-of-the-art worst-case regret bounds and substantially tighter bounds if the RL environment has special features but without apriori knowledge of the environment from the algorithm.

735, TITLE: Global Convergence of Block Coordinate Descent in Deep Learning

<http://proceedings.mlr.press/v97/zeng19a.html>

AUTHORS: Jinshan Zeng, Tim Tsz-Kit Lau, Shaobo Lin, Yuan Yao

HIGHLIGHT: In this paper, we aim at providing a general methodology for provable convergence guarantees for this type of methods.

736, TITLE: Making Convolutional Networks Shift-Invariant Again

<http://proceedings.mlr.press/v97/zhang19a.html>

- AUTHORS: Richard Zhang  
HIGHLIGHT: We show that when integrated correctly, it is compatible with existing architectural components, such as max-pooling.
- 737, TITLE: Warm-starting Contextual Bandits: Robustly Combining Supervised and Bandit Feedback  
<http://proceedings.mlr.press/v97/zhang19b.html>  
AUTHORS: Chicheng Zhang, Alekh Agarwal, Hal Daum? Iii, John Langford, Sahand Negahban  
HIGHLIGHT: We investigate the feasibility of learning from both fully-labeled supervised data and contextual bandit data.
- 738, TITLE: When Samples Are Strategically Selected  
<http://proceedings.mlr.press/v97/zhang19c.html>  
AUTHORS: Hanrui Zhang, Yu Cheng, Vincent Conitzer  
HIGHLIGHT: In this paper, we introduce a theoretical framework for this problem and provide key structural and computational results.
- 739, TITLE: Self-Attention Generative Adversarial Networks  
<http://proceedings.mlr.press/v97/zhang19d.html>  
AUTHORS: Han Zhang, Ian Goodfellow, Dimitris Metaxas, Augustus Odena  
HIGHLIGHT: In this paper, we propose the Self-Attention Generative Adversarial Network (SAGAN) which allows attention-driven, long-range dependency modeling for image generation tasks.
- 740, TITLE: Circuit-GNN: Graph Neural Networks for Distributed Circuit Design  
<http://proceedings.mlr.press/v97/zhang19e.html>  
AUTHORS: Guo Zhang, Hao He, Dina Katabi  
HIGHLIGHT: We present Circuit-GNN, a graph neural network (GNN) model for designing distributed circuits.
- 741, TITLE: LatentGNN: Learning Efficient Non-local Relations for Visual Recognition  
<http://proceedings.mlr.press/v97/zhang19f.html>  
AUTHORS: Songyang Zhang, Xuming He, Shipeng Yan  
HIGHLIGHT: In this work, we propose an efficient and yet flexible non-local relation representation based on a novel class of graph neural networks.
- 742, TITLE: Neural Collaborative Subspace Clustering  
<http://proceedings.mlr.press/v97/zhang19g.html>  
AUTHORS: Tong Zhang, Pan Ji, Mehrtash Harandi, Wenbing Huang, Hongdong Li  
HIGHLIGHT: We introduce the Neural Collaborative Subspace Clustering, a neural model that discovers clusters of data points drawn from a union of low-dimensional subspaces.
- 743, TITLE: Incremental Randomized Sketching for Online Kernel Learning  
<http://proceedings.mlr.press/v97/zhang19h.html>  
AUTHORS: Xiao Zhang, Shizhong Liao  
HIGHLIGHT: To address these issues, we propose a novel incremental randomized sketching approach for online kernel learning, which has efficient incremental maintenances with theoretical guarantees.
- 744, TITLE: Bridging Theory and Algorithm for Domain Adaptation  
<http://proceedings.mlr.press/v97/zhang19i.html>  
AUTHORS: Yuchen Zhang, Tianle Liu, Mingsheng Long, Michael Jordan  
HIGHLIGHT: We introduce Margin Disparity Discrepancy, a novel measurement with rigorous generalization bounds, tailored to the distribution comparison with the asymmetric margin loss, and to the minimax optimization for easier training.
- 745, TITLE: Adaptive Regret of Convex and Smooth Functions



<http://proceedings.mlr.press/v97/zhang19j.html>

AUTHORS: Lijun Zhang, Tie-Yan Liu, Zhi-Hua Zhou

HIGHLIGHT: To this end, we develop novel adaptive algorithms for convex and smooth functions, and establish problem-dependent regret bounds over any interval.

746, TITLE: Random Function Priors for Correlation Modeling

<http://proceedings.mlr.press/v97/zhang19k.html>

AUTHORS: Aonan Zhang, John Paisley

HIGHLIGHT: In this paper, we introduce random function priors for  $Z_n$  for modeling correlations among its  $K$  dimensions  $Z_{n1}$  through  $Z_{nK}$ , which we call population random measure embedding (PRME).

747, TITLE: Co-Representation Network for Generalized Zero-Shot Learning

<http://proceedings.mlr.press/v97/zhang19l.html>

AUTHORS: Fei Zhang, Guangming Shi

HIGHLIGHT: Hence we propose a embedding model called co-representation network to learn a more uniform visual embedding space that effectively alleviates the bias problem and helps with classification.

748, TITLE: SOLAR: Deep Structured Representations for Model-Based Reinforcement Learning

<http://proceedings.mlr.press/v97/zhang19m.html>

AUTHORS: Marvin Zhang, Sharad Vikram, Laura Smith, Pieter Abbeel, Matthew Johnson, Sergey Levine

HIGHLIGHT: In this paper, we present a method for learning representations that are suitable for iterative model-based policy improvement, even when the underlying dynamical system has complex dynamics and image observations, in that these representations are optimized for inferring simple dynamics and cost models given data from the current policy.

749, TITLE: A Composite Randomized Incremental Gradient Method

<http://proceedings.mlr.press/v97/zhang19n.html>

AUTHORS: Junyu Zhang, Lin Xiao

HIGHLIGHT: We propose a composite randomized incremental gradient method by extending the SAGA framework.

750, TITLE: Fast and Stable Maximum Likelihood Estimation for Incomplete Multinomial Models

<http://proceedings.mlr.press/v97/zhang19o.html>

AUTHORS: Chenyang Zhang, Guosheng Yin

HIGHLIGHT: We propose a fixed-point iteration approach to the maximum likelihood estimation for the incomplete multinomial model, which provides a unified framework for ranking data analysis.

751, TITLE: Theoretically Principled Trade-off between Robustness and Accuracy

<http://proceedings.mlr.press/v97/zhang19p.html>

AUTHORS: Hongyang Zhang, Yaodong Yu, Jiantao Jiao, Eric Xing, Laurent El Ghaoui, Michael Jordan

HIGHLIGHT: In this work, we decompose the prediction error for adversarial examples (robust error) as the sum of the natural (classification) error and boundary error, and provide a differentiable upper bound using the theory of classification-calibrated loss, which is shown to be the tightest possible upper bound uniform over all probability distributions and measurable predictors.

752, TITLE: Learning Novel Policies For Tasks

<http://proceedings.mlr.press/v97/zhang19q.html>

AUTHORS: Yunbo Zhang, Wenhao Yu, Greg Turk

HIGHLIGHT: In this work, we present a reinforcement learning algorithm that can find a variety of policies (novel policies) for a task that is given by a task reward function.

753, TITLE: Greedy Orthogonal Pivoting Algorithm for Non-Negative Matrix Factorization

<http://proceedings.mlr.press/v97/zhang19r.html>

AUTHORS: Kai Zhang, Sheng Zhang, Jun Liu, Jun Wang, Jie Zhang

**HIGHLIGHT:** To address this challenge, we propose an innovative procedure called Greedy Orthogonal Pivoting Algorithm (GOPA).

754, **TITLE:** Interpreting Adversarially Trained Convolutional Neural Networks  
<http://proceedings.mlr.press/v97/zhang19s.html>

**AUTHORS:** Tianyuan Zhang, Zhanxing Zhu

**HIGHLIGHT:** We design systematic approaches to interpret AT-CNNs in both qualitative and quantitative ways and compare them with normally trained models.

Second, to achieve quantitative verification, we construct additional test datasets that destroy either textures or shapes, such as style-transferred version of clean data, saturated images and patch-shuffled ones, and then evaluate the classification accuracy of AT-CNNs and normal CNNs on these datasets.

755, **TITLE:** Adaptive Monte Carlo Multiple Testing via Multi-Armed Bandits

<http://proceedings.mlr.press/v97/zhang19t.html>

**AUTHORS:** Martin Zhang, James Zou, David Tse

**HIGHLIGHT:** In this paper, we propose \texttt{A}daptive \texttt{M}C multiple \texttt{T}esting (\texttt{AMT}) to estimate MC p-values and control false discovery rate in multiple testing.

756, **TITLE:** On Learning Invariant Representations for Domain Adaptation

<http://proceedings.mlr.press/v97/zhao19a.html>

**AUTHORS:** Han Zhao, Remi Tachet Des Combes, Kun Zhang, Geoffrey Gordon

**HIGHLIGHT:** To give a sufficient condition for domain adaptation, we propose a natural and interpretable generalization upper bound that explicitly takes into account the aforementioned shift.

757, **TITLE:** Metric-Optimized Example Weights

<http://proceedings.mlr.press/v97/zhao19b.html>

**AUTHORS:** Sen Zhao, Mahdi Milani Fard, Harikrishna Narasimhan, Maya Gupta

**HIGHLIGHT:** Motivated by known connections between complex test metrics and cost-weighted learning, we propose addressing these issues by using a weighted loss function with a standard loss, where the weights on the training examples are learned to optimize the test metric on a validation set.

758, **TITLE:** Improving Neural Network Quantization without Retraining using Outlier Channel Splitting

<http://proceedings.mlr.press/v97/zhao19c.html>

**AUTHORS:** Ritchie Zhao, Yuwei Hu, Jordan Dotzel, Chris De Sa, Zhiru Zhang

**HIGHLIGHT:** In this work, we propose outlier channel splitting (OCS), which duplicates channels containing outliers, then halves the channel values.

759, **TITLE:** Maximum Entropy-Regularized Multi-Goal Reinforcement Learning

<http://proceedings.mlr.press/v97/zhao19d.html>

**AUTHORS:** Rui Zhao, Xudong Sun, Volker Tresp

**HIGHLIGHT:** On a set of multi-goal robotic tasks of OpenAI Gym, we compare our method with other baselines and show promising improvements in both performance and sample-efficiency.

760, **TITLE:** Stochastic Iterative Hard Thresholding for Graph-structured Sparsity Optimization

<http://proceedings.mlr.press/v97/zhou19a.html>

**AUTHORS:** Baojian Zhou, Feng Chen, Yiming Ying

**HIGHLIGHT:** In this paper, we propose a stochastic gradient-based method for solving graph-structured sparsity constraint problems, not restricted to the least square loss.

761, **TITLE:** Lower Bounds for Smooth Nonconvex Finite-Sum Optimization

<http://proceedings.mlr.press/v97/zhou19b.html>

**AUTHORS:** Dongruo Zhou, Quanquan Gu

**HIGHLIGHT:** In this paper, we study the lower bounds for smooth nonconvex finite-sum optimization, where the objective function is the average of  $n$  nonconvex component functions.

- 762, TITLE: Lipschitz Generative Adversarial Nets  
<http://proceedings.mlr.press/v97/zhou19c.html>  
AUTHORS: Zhiming Zhou, Jiadong Liang, Yuxuan Song, Lantao Yu, Hongwei Wang, Weinan Zhang, Yong Yu, Zhihua Zhang  
HIGHLIGHT: In this paper we show that generative adversarial networks (GANs) without restriction on the discriminative function space commonly suffer from the problem that the gradient produced by the discriminator is uninformative to guide the generator.
- 763, TITLE: Toward Understanding the Importance of Noise in Training Neural Networks  
<http://proceedings.mlr.press/v97/zhou19d.html>  
AUTHORS: Mo Zhou, Tianyi Liu, Yan Li, Dachao Lin, Enlu Zhou, Tuo Zhao  
HIGHLIGHT: This implies that the noise enables the algorithm to efficiently escape from the spurious local optimum.
- 764, TITLE: BayesNAS: A Bayesian Approach for Neural Architecture Search  
<http://proceedings.mlr.press/v97/zhou19e.html>  
AUTHORS: Hongpeng Zhou, Minghao Yang, Jun Wang, Wei Pan  
HIGHLIGHT: In this paper, we employ the classic Bayesian learning approach to alleviate these two issues by modeling architecture parameters using hierarchical automatic relevance determination (HARD) priors.
- 765, TITLE: Transferable Clean-Label Poisoning Attacks on Deep Neural Nets  
<http://proceedings.mlr.press/v97/zhu19a.html>  
AUTHORS: Chen Zhu, W. Ronny Huang, Hengduo Li, Gavin Taylor, Christoph Studer, Tom Goldstein  
HIGHLIGHT: In this paper, we explore clean-label poisoning attacks on deep convolutional networks with access to neither the network's output nor its architecture or parameters.
- 766, TITLE: Improved Dynamic Graph Learning through Fault-Tolerant Sparsification  
<http://proceedings.mlr.press/v97/zhu19b.html>  
AUTHORS: Chunjiang Zhu, Sabine Storandt, Kam-Yiu Lam, Song Han, Jinbo Bi  
HIGHLIGHT: We propose a new type of graph sparsification namely fault-tolerant (FT) sparsification to significantly reduce the cost to only a constant.
- 767, TITLE: Poission Subsampled R?nyi Differential Privacy  
<http://proceedings.mlr.press/v97/zhu19c.html>  
AUTHORS: Yuqing Zhu, Yu-Xiang Wang  
HIGHLIGHT: We consider the problem of privacy-amplification by under the Renyi Differential Privacy framework.
- 768, TITLE: Learning Classifiers for Target Domain with Limited or No Labels  
<http://proceedings.mlr.press/v97/zhu19d.html>  
AUTHORS: Pengkai Zhu, Hanxiao Wang, Venkatesh Saligrama  
HIGHLIGHT: We propose a novel visual attribute encoding method that encodes each image as a low-dimensional probability vector composed of prototypical part-type probabilities.
- 769, TITLE: The Anisotropic Noise in Stochastic Gradient Descent: Its Behavior of Escaping from Sharp Minima and Regularization Effects  
<http://proceedings.mlr.press/v97/zhu19e.html>  
AUTHORS: Zhanxing Zhu, Jingfeng Wu, Bing Yu, Lei Wu, Jinwen Ma  
HIGHLIGHT: Along this line, we study a general form of gradient based optimization dynamics with unbiased noise, which unifies SGD and standard Langevin dynamics.
- 770, TITLE: Surrogate Losses for Online Learning of Stepsizes in Stochastic Non-Convex Optimization

<http://proceedings.mlr.press/v97/zhuang19a.html>

AUTHORS: Zhenxun Zhuang, Ashok Cutkosky, Francesco Orabona

HIGHLIGHT: In this paper, we propose new surrogate losses to cast the problem of learning the optimal stepsizes for the stochastic optimization of a non-convex smooth objective function onto an online convex optimization problem.

771, TITLE: Latent Normalizing Flows for Discrete Sequences

<http://proceedings.mlr.press/v97/ziegler19a.html>

AUTHORS: Zachary Ziegler, Alexander Rush

HIGHLIGHT: We propose a VAE-based generative model which jointly learns a normalizing flow-based distribution in the latent space and a stochastic mapping to an observed discrete space.

772, TITLE: Beating Stochastic and Adversarial Semi-bandits Optimally and Simultaneously

<http://proceedings.mlr.press/v97/zimmert19a.html>

AUTHORS: Julian Zimmert, Haipeng Luo, Chen-Yu Wei

HIGHLIGHT: We develop the first general semi-bandit algorithm that simultaneously achieves  $\mathcal{O}(\log T)$  regret for stochastic environments and  $\mathcal{O}(\sqrt{T})$  regret for adversarial environments without knowledge of the regime or the number of rounds  $T$ .

773, TITLE: Fast Context Adaptation via Meta-Learning

<http://proceedings.mlr.press/v97/zintgraf19a.html>

AUTHORS: Luisa Zintgraf, Kyriacos Shiarli, Vitaly Kurin, Katja Hofmann, Shimon Whiteson

HIGHLIGHT: We propose CAVIA for meta-learning, a simple extension to MAML that is less prone to meta-overfitting, easier to parallelise, and more interpretable.

774, TITLE: Natural Analysts in Adaptive Data Analysis

<http://proceedings.mlr.press/v97/zrnic19a.html>

AUTHORS: Tijana Zrnic, Moritz Hardt

HIGHLIGHT: In this work, we propose notions of natural analysts that smoothly interpolate between the optimal non-adaptive bounds and the best-known adaptive generalization bounds.